Studies on the effects of depletion of the chaperonin GroEL in *Escherichia coli.*

John Acord

A thesis presented for the degree of Ph.D.

Institute of Cell and Molecular Biology
University of Edinburgh

September 2001

Acknowledgements

To Mr. G. Thank you for standing by me through the good times and the bad. I wouldn't have made it without you.

To my mum. Thanks for your unending support (financial and otherwise!) I wouldn't have made it without you either!

To Dawn. You set the standard by which all coffee breaks will be judged.

And finally to Millie. For putting up with a temperamental Englishman and never getting too upset!

Abstract

The GroE heat shock proteins (GroEL and GroES) of *Escherichia coli* represent major molecular chaperones that participate in folding and assembly of a variety of proteins and are essential for cell growth at all temperatures. From *in vitro* studies, GroEL is thought to be highly promiscuous in substrate binding, interacting with almost any non-native model protein. However, *in vivo*, GroEL is involved in the folding of only 10-15% of newly translated polypeptides, suggesting specificity for a defined set of substrates.

In an attempt to identify GroE substrates, a strain of *E.coli* has been constructed in our laboratory in which expression of GroE can be turned off. Previous work from our laboratory had suggested that DapA, the first enzyme on the lysine biosynthetic pathway, was affected by GroE depletion in *Escherichia coli*. Native DapA is present as a homotetramer; in this study, work is described that suggests that DapA achieves a homotetrameric state in GroE depleted cells but lacks enzymatic activity. Work is also described that suggests that native DapA cannot be overproduced in GroE depleted cells as it becomes highly susceptible to aggregation.

Work in chapter 4 suggests that the *lac* inhibitor protein, LacI, is also affected by GroE depletion. *In vitro* footprinting studies in GroE depleted cells suggest that LacI remains bound to the *lac* operator sequence in the presence of the gratuitous inducer IPTG.

Previous work in other laboratories had suggested that *dapA* is not regulated at the level of expression. In this work data is presented that suggests that *dapA* is regulated at the level of expression, possibly by a transcriptional activator, in response to intracellular levels of the molecule diaminopimelic acid.

Abbreviations

ADP	Adenosine 5'-diphosphate
AMP	Adensonie 5'-monophosphate
Amp	Ampicillin
APS	Ammonium persulphate
ara	Arabinose
ATP	Adenosine 5'-triphosphate
bp	Base pair
BSA	Bovine serum albumin
cAMP	Cyclic AMP
CAP	Catabolite gene activating protein
Cmp	Chloramphenicol
CTP	Cytosine 5'-triphosphate
Da	Dalton
DAP	Diaminopimelic acid
dATP	Deoxy ATP
ddATP	Dideoxy ATP
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EC	Enzyme classification
EDTA	Diaminoethanetetraacetic acid
EtBr	Ethidium bromide
glu	Glucose
GTP	Guanosine 5'-triphosphate
IPTG	Isopropyl-β-D-thiogalactoside
Kan	Kanamycin
kb	Kilobase
kDa	Kilodalton
mRNA	Messenger RNA
OD_w	Optical density $_{\text{wavelength}}$
ONPG	O-nitrophenyl-β-D-galactoside
PAGE	Polyacrylamide gel electrophoresis

PCR	Polymerase chain reaction
r.p.m.	Revolutions per minute
rbs	Ribosome binding site
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulphate
Spc	Spectinomycin
ssDNA	Single stranded DNA
Str	Streptomycin
TAE	Tris-acetate EDTA
TE	Tris-EDTA
TEMED	N,N,N',N',-tetramethylethylethylenediamine
Tet	Tetracyclin
Tris	2-amino-2-hydroxymethyl-1,3-propandiol
^{ts}	Temperature sensitive
TTP	Thymidine 5'-triphosphate
UTP	Uridine 5'-triphosphate
v:v	Volume per volume
w:v	Weight per volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
X^R	Resistance to antibiotic X
X^S	Sensitive to antibiotic X

Chapter I

Introduction

1.1 The protein folding problem

Protein folding is arguably the single most important process in biology because it converts linear polypeptide chains into 3-dimensional structures that endow proteins with their activities. Traditionally the protein-folding problem has been studied *in vitro* with purified, chemically denatured polypeptides as substrates. The pioneering studies of Anfinsen and his colleagues (Anfinsen, 1973) with the *in vitro* refolding of purified ribonuclease A left the long-lasting impression that the folding of a newly synthesised polypeptide was an intrinsic feature of its primary structure, independent of other factors (Jaenicke, 1991). The wealth of *in vitro* studies that followed these key experiments has indeed shaped our understanding of the general kinetic and thermodynamic principles of protein folding, at least for small, globular proteins (Baldwin, 1994; Matthews, 1993; Ptitsyn, 1995). Briefly, within a few milliseconds of the initiation of folding, the extended polypeptide chain collapses, driven primarily by hydrophobic forces, to a compact folding intermediate. This intermediate, named the molten globule, is stabilised by flexible interactions between hydrophobic residues in the core of the molecule and possesses elements of secondary structure, such as α-helices and β-sheets, but lacks the ordered interactions between these elements which form the tertiary structure in the native state (Christensen et al, 1991; Creighton, 1990; Radford et al, 1992). This leads to the exposure of hydrophobic protein surfaces to solvent, which makes this intermediate prone to aggregation, an off pathway folding reaction that is considered irreversible under most conditions. Folding from the molten globule intermediate towards the native state is a comparatively slow process that occurs on the time scale of seconds to minutes. As the polypeptide progresses along this pathway, it is believed to pass through various incompletely folded intermediates via a series of structural rearrangements.

The probability that a given unfolded polypeptide will fold properly increases at relatively low protein concentration (which limits interpolypeptide aggregation) and low temperature (which attenuates hydrophobic interaction). However the relatively high protein concentration in the cytososl (300 mg ml^{-1}) subjects the growing polypeptide to premature interactions with other intra- or interpolypeptide domains, thereby leading to misfolding and aggregation (Jaenicke, 1991). To deal with these problems, a set of proteins, collectively called chaperones, has evolved, their primary function is to ensure that polypeptides will fold properly in the cell. These chaperone proteins act primarily by binding to the reactive surfaces (mostly hydrophobic) of polypeptides. In doing so, chaperones sequester these reactive sites

and prevent interactions with other reactive sites in their vicinity, effectively preventing aggregation and favouring the proper folding pathway. In general the chaperone proteins act without covalently modifying their polypeptide substrates and without being part of the finished product (Ellis et al, 1991). Because high temperature tend to favour protein unfolding and hydrophobic interactions, there is an extra need for chaperones to prevent protein aggregation *in vivo*. This is most likely the reason why many chaperones are expressed at much higher levels in cells after heat shock (Georgopoulos et al, 1993).

1.2 The *E.coli* GroE chaperonin machine

1.2.1 Discovery of the groE *locus*

In the 1970s, a number of laboratories reported the isolation of *E.coli* mutants that block growth of either bacteriophage λ or T4, specifically at the level of head assembly, and bacteriophage T5 at the level of tail assembly (Table 1.2.1). These mutants were variously termed *groE* (Georgopoulos et al, 1973; Georgopoulos et al, 1972; Sternberg, 1973a; Sternberg, 1973b), *tabB* (Coppo et al, 1973; Takahashi et al, 1975), *mop* (Takano et al, 1972), *hdh* (Revel et al, 1980), and *hdh* (*fatA*) (Simon et al, 1974). The *groE*, *mop*, *tabB,* and *hdh* alleles turned out to be mutations at the same locus and are referred to here collectively as *groE*. The *groE* mutants were so named because they blocked bacteriophage λ growth (*gro*), and the first bacteriophage λ mutants able to propagate on these mutant bacteria were shown to have amber mutations in gene *E*, whose product is the major capsid protein of bacteriophage λ, hence the designation *groE* (Georgopoulos et al, 1973; Sternberg, 1973b). The mechanism by which these mutations in gene *E* bypassed the *groE* imposed block was by lowering the amount of the λ*E* gene product, thus enabling the mutated GroE to concentrate on fewer bacteriophage capsid structures (Georgopoulos et al, 1973; Sternberg, 1973a)

The molecular characterisation of the *groE* locus was aided by the isolation of λ*groE*$^+$ transducing bacteriophages. The λ*groE*$^+$ transducing bacteriophages were selected directly out of a pool of λ transducing bacteriophages, each carrying a different segment of the *E. coli* chromosome, simply based on their ability to form plaques on various *groE* mutant hosts (Georgopoulos et al, 1978; Hendrix et al, 1978). These λ*groE*$^+$ transducing bacteriophages also served to produce high levels of GroE protein, thus leading not only to the preliminary identification of the GroE product as a 60 kDa protein, but also to its purification (Hendrix, 1979; Hohn et al, 1979). However, further deletion and point mutation analysis of the various λ*groE*$^+$ transducing bacteriophages demonstrated that the *groE* locus is actually made up of two genes, *groES*, encoding a protein with an apparent mass of 15 kDa, and *groEL*, encoding the previously identified 60 kDa protein (Tilly et al, 1981). Interestingly, the isolated *groE* mutations were approximately equally distributed between the *groES* and *groEL* genes. Since mutations in either *groES* or *groEL* affected bacteriophage λ and Mu head morphogenesis, Mu and T5 tail morphogenesis (Georgopoulos et al, 1973; Sternberg, 1973b; Grimaud et al, 1998; Zweig et al,

1973), as well as bacterial growth at high temperature, qualitatively and quantitatively to the same extent, the GroES and GroEL proteins must function in the same biochemical pathway. Furthermore, the realisation that the Rubisco binding protein of plants, necessary for the proper assembly of Rubisco, is homologous to GroEL exemplified their universal sequence conservation, as well as their function in promoting oligomeric protein assembly (Hemmingsen et al, 1988).

It took a long time before it was realised that the GroES and GroEL proteins constitute a molecular chaperone machine. First the original *groE* alleles were isolated on the basis of blocking bacteriophage λ or T4 growth, and the fact that they defined two distinct but adjacent genes was not discovered until 10 years later. Because the *groE* mutation blocked the head protein morphogenic cleavages intrinsic to λ and T4 (Georgopoulos et al, 1973; Grimaud et al, 1998), and the tail morphogenic cleavage associated with T5 tail assembly (Zweig et al, 1973), this left the impression that the GroE protein itself might be a protease or a component of a protease. The purification and characterisation of GroE and the realisation that it is a tetradecameric structure with sevenfold symmetry encouraged the notion that its symmetric surface acts like a scaffold to allow the proper oligomerisation of some structural proteins to give rise to the correct morphogenic structure, which in turn is the correct substrate for protease cleavage.

Host Mutation	λ	Mu	T4	T5
groES	Head assembly (genes *B* and *E*)	Head assembly (gene *H*); tail assembly	No effect	Tail assembly (gene *D19*)
groEL	Head assembly (genes *B* and *E*)	Head assembly (gene *X*); tail assembly	Head assembly (genes *31* and *23*)	Tail assembly (gene *D19*)

Table 1.2.1 Effects of *E.coli groE* mutations on bacteriophage growth. The various *groE* mutations block either head assembly and/or tail assembly. Also shown, in parentheses, are the bacteriophage-encoded genes in which mutations can occur, thus permitting growth on the *groE* mutants. Adapted from Ang *et al.* which should also be consulted for individual references (Ang et al, 2000).

1.2.2 The GroE proteins are molecular chaperones

The term 'molecular chaperone' was first used to describe the role of nucleoplasmin, a protein involved with assembling DNA and histones isolated from *Xenopus* oocytes into ordered nucleosomes (Laskey et al, 1980). Without the nucleoplasmin the DNA and histones formed an unorganised aggregate *in vitro*. Ellis adopted the term to describe a protein associated with Rubisco (ribulose-1-5-bisphosphate carboxylase oxygenase), an enzyme involved in fixing CO_2 in higher plant chloroplasts. Rubisco is a multimer of eight small and eight large subunits and can amount to as much as 50% of a plant's total cell protein. The large subunits are encoded by plastid genes and the small subunits by nuclear genes, and are thus imported from the cytoplasm. For attainment of the correct Rubisco conformation, another protein called the Rubisco subunit binding protein (RSBP) is required. RSBP does not form a part of the final Rubisco enzyme, but is required during the assembly of the large subunit octamer from four dimers (Barraclough et al, 1980). Ellis called his RSBP a molecular chaperone and showed that it was made up of two types of subunit (α and β) found in equal amounts. The α and β subunits showed 50% sequence similarity to each other, and interestingly RBSP shares 50% sequence similarity with GroEL (Hemmingsen et al, 1988). RSBP was shown to have the same tetradecameric ring structure and sevenfold symmetry as GroEL.

The original definition of a molecular chaperone was a protein that brings potentially complementary surfaces together, but makes sure that no improper or premature interactions occur. Ellis' adoption of the term 'molecular chaperone' to explain the action of RSBP has now been extended to cover a growing list of different proteins that bind polypeptides and mediate their folding and/or assembly. Ellis proposed that molecular chaperones of the GroEL/HSP-60 type (which are required for cell viability at all temperatures) should be designated the chaperonins, which seemed more appropriate than terminology that implies their involvement solely during heat-shock (Ellis et al, 1991; Ellis, 1993).

It is interesting to note that while many bacteria species are reported to have only one copy of the *groEL/hsp60* gene, a number are reported to have two or more copies. Bacteria which exhibit multiple copies of the *groEL/hsp60* gene include several *Streptomyces* species, *Mycobacterium tuberculosis*, *M. leprae*, *Rhizobium meliloti*, *Bradyrhizobium japonicum, Synechocystis* sp. PCC 6803 and *Rhizobium leguminosarum* (Guglielmi et al, 1991; De Wit et al, 1992; Kong et al, 1993; Rusanganwa et al, 1993; Fischer et al, 1993; Lehel et al, 1993; Wallington et al, 1994). While the functional significance of these multiple copies remains unknown,

it has been suggested that multiple Hsp60 proteins may serve different functions within the bacteria (Wallington et al, 1994).

1.2.3 In vitro studies of GroEL

The realisation that GroEL acts solely at the level of preventing the aggregation and aiding the correct folding of a single polypeptide chain had to await subsequent work from many laboratories. The first *in vitro* demonstration of polypeptide binding by GroEL was by Bochkareva *et al* (Bochkareva et al, 1988). They showed that *E. coli* GroEL protein bound newly translated unfolded pre-β-lactamase (pre-β-lac) and chloramphenicol acetyl transferase (CAT), both of which need to be in an unfolded conformation in order to traverse the bacterial plasma membrane into the periplasmic space. A transcription/translation system was used to produce the pre-β-lac and CAT polypeptides; an N-acyl-Met-$tRNA^{met}$ with the acyl group containing a precursor of the highly photoreactive carbene radical was used to add the initial N-terminal methionine residue, resulting in polypeptides that are labelled at one site only. The *E. coli* translation extract (S30) contained GroEL, but this could be removed (up to ~ 95%) by passing the extract through an anti-GroEL sephadex affinity column. Exogenous GroEL was then added to the translation mix. Photoactivation led to the pre-β-lac or CAT polypeptides binding to the GroEL tetradecamers. Furthermore, it was shown that the protein was binding only when it was in an unfolded state. Addition of Mg-ATP, but not a non-hydrolysable analogue, to the GroEL-peptide complex resulted in the dissociation of the polypeptide, indicating a role for GroEL's previously observed weak ATPase activity.

A more detailed study of the interaction between pre-β-lac and GroEL was conducted by Zahn and Plückthun (Zahn et al, 1992). They showed that an early folding intermediate of pre-β-lac is recognised by GroEL and thus protected from aggregation. The form that is discharged from GroEL by Mg-ATP hydrolysis is another folding intermediate, which can misfold to a non-native conformation (but does not form an aggregate) in a pH-dependant fashion. This intermediate is not bound to GroEL and would fold correctly under normal physiological conditions without any external assistance. They therefore suggest that GroEL prevents aggregation by binding polypeptides that have the capacity to aggregate, and processing the polypeptides to conformations that will not aggregate. Interestingly they also showed that mature-β-lac does not bind to GroEL under any folding

conditions and so proposed that the N-terminal signal sequence is a motif recognised by GroEL.

The binding of unfolded polypeptides by GroEL seems consistent with its proposed role in the heat-shocked cell, in that thermally denatured proteins are being sequestered and thus prevented from aggregating. However, the work of Bochkareva (Bochkareva et al, 1988) also suggests a role for GroEL in membrane translocation of proteins, and indeed Kusukawa (Kusukawa et al, 1989) has shown that the export of pre-β-lac is defective in *groE*^{ts} mutants *in vivo*, although protein export is not thought to be a major function of the GroE proteins as other specialised export systems are normally employed (Kumamoto, 1991).

Lissin (Lissin et al, 1990) examined the assembly of the GroEL tetradecamer itself *in vitro*. It was shown that GroEL tetradecamers could be formed from GroEL monomers in the presence of ATP and Mg^{2+} ions at 23°C. Assembly intermediates were not found, suggesting that the assembly of the GroEL tetradecamer is highly cooperative. Assembly of the GroEL monomers was also found to be enhanced by the presence of GroEL tetradecamers. The authors refer to this as 'self-chaperoning'.

An attempt to determine how many *E. coli* proteins could bind to GroEL was performed by Viitanin (Viitanen et al, 1992). Here they used a soluble protein cell extract labelled with ^{35}S-methionine and denatured in 5 M guanidine-HCl, to show that about half of the unfolded protein complex could be bound to GroEL tetradecamers and subsequently released on addition of Mg-ATP. This rather crude analysis showed that a wide variety of denatured *E. coli* proteins have the potential to bind to GroEL, but did not attempt to address the question of which proteins are bound and which are not.

Landry and Gierasch (Landry et al, 1991a; Landry et al, 1991b) observed that the experimental *in vitro* GroEL substrates Rubisco, β-lactamase, citrate synthase and rhodanese all contain solvent-exposed amphipathic amino-terminal α-helices and suggest that it was this secondary structure that was recognised by GroEL; others (Brazil et al, 1997; Kobayashi et al, 1999) have also obtained similar results. However, Schmidt and Buchner (Schmidt et al, 1992) reported that a protein fragment consisting entirely of β-sheet secondary structure elements is bound and released by the GroEL-ES-ATP system *in vitro* [also reported by Chen (Chen et al, 1999)]. While it is now generally accepted that GroEL tetradecamers interact with polypeptides that are essentially unfolded but still contain a reasonable amount of secondary structure (Martin et al, 1991) no sequence specificity has been found in substrates, and the nature of the binding interactions has only recently been addressed in stereochemical terms (Chatellier et al, 1999; Kobayashi et al, 1999;

Tanaka et al, 1999). It is also accepted that exposed hydrophobic surfaces on substrates are primarily responsible for recognition by GroEL (Chatellier et al, 1999; Itzhaki et al, 1995; Lin et al, 1995), consistent with the ability of GroEL to recognise the exposed hydrophobic surfaces that are characteristic of non-native proteins. In addition GroEL prefers substrates bearing net positive charges (Chatellier et al, 1999; Dessauer et al, 1994; Itzhaki et al, 1995; Rosenberg et al, 1993; Koll et al, 1992), indicating an electrostatic component in GroEL-substrate interactions (Perrett et al, 1997).

1.2.4 In vivo studies of GroEL

Demonstration of the formation of complexes between certain proteins and GroEL in *E. coli* has been reported by several groups (Carrillo et al, 1992; Dolan et al, 1992; Lee et al, 1992; Dionisi et al, 1998). All of these examples have used proteins foreign to *E. coli*, but have shown that the protein of interest would not be functional and/or assembled in *groE*-mutant bacteria. More interestingly it was shown that proteins that normally form insoluble aggregates when overexpressed in *E. coli* can be largely solubilised by concomitant overexpression of the *groE* genes. These results suggest that the GroE proteins are not highly selective in their choice of substrate and will help foreign proteins to assemble correctly. This fact has been used to obtain functional foreign proteins from genes cloned in *E. coli* (King et al, 1996; Hahm et al, 2001; Goenka et al, 2001).

From *in vitro* work it was assumed that GroEL is highly promiscuous and will interact with any protein (Coyle et al, 1997) that can be accommodated in the cis-cavity (Sakikawa et al, 1999). But calculations have suggested that there is only enough GroEL present in the cell to fold 5-10% (Lorimer, 1996) or 10-15% (Ewalt et al, 1997) of all newly synthesised proteins; thus suggesting that GroEL may have specificity for a defined set of substrates. There have been a number of attempts to identify GroE substrates *in vivo* in *E. coli*. Horwich (Horwich et al, 1993) made a *groEL* mutant that was severely temperature sensitive, and therefore unable to form colonies above 35°C. Using this mutant he was able to demonstrate the loss of solubility of 16 diverse polypeptides on 2D protein gels, of which three were identified and found to be citrate synthase, ketogluterate dehydrogenase and polynucleotide phosphorylase. Overall, 30% of all the polypeptides were affected by the loss of GroEL function in the mutant cells at restrictive temperatures. These results suggest that there is a specific set of proteins that rely on GroEL to allow their

folding to the native conformation. Since, in this study, extracts were prepared from cells containing substantial quantities of inactive mutant GroEL, it is possible that the reduced solubilities observed result from interaction with inactivated GroEL molecules, rather than because of failure of assisted folding.

Kanemori (Kanemori et al, 1994) attempted to identify GroE substrates by reducing the levels of GroE in *E. coli*. A strain was constructed in which the GroE protein level was manipulated by placing *groE* under the control of the *lacUV5* promoter on a multicopy plasmid in a strain lacking the chromosomal *groE* operon. When the strain was grown with a limited concentration of inducer (IPTG) at 37°C, the GroE level and growth rate were comparable to those of the wild type. When the cells were depleted of IPTG, they continued to grow at or below 37°C albeit at reduced rates, despite the reduced levels of GroE (~ 25% of wild type). Under these conditions the cellular contents of at least 13 polypeptides were affected. Among the most striking effects was the enhanced synthesis of a set of heat shock proteins, which resulted from the increased level of σ^{32}, which is required for the transcription of heat shock genes. Other proteins affected by the reduced GroE level included two proteins [enzymes of the Entner-Doudoroff pathway (Egan et al, 1992)] encoded by the *edd-eda* operon and the ribosomal protein S6, suggesting that the GroE chaperones are involved in regulating expression of genes for carbohydrate metabolism and in modulating biogenesis or function of ribosomes. Note that in this study GroE levels were only reduced to 25% of the wild type level, a far cry from complete GroE depletion.

Ewalt (Ewalt et al, 1997) performed experiments to determine the polypeptide flux through the bacterial chaperonin system. To determine the extent to which GroEL interacts with newly synthesised proteins, GroEL and its bound substrates were isolated by immunoprecipitation from pulse labelled *E. coli*. Cells were labelled at temperatures between 30°C and 42°C for 15 seconds with ^{35}S-methionine followed by the addition of excess nonradioactive methionine. At specific times, cells were removed, rapidly cooled to 4°C, converted to spheroplasts, and lysed in EDTA-containing buffer [low temperature and chelating Mg^{2+} stop ATP-dependant polypeptide release from GroEL (Martin et al, 1991)]. The data showed that a diverse set of newly synthesised polypeptides, predominantly between 10-55 kDa, interacts with GroEL, accounting for 10-15% of all cytoplasmic protein under normal growth conditions, and for 30% or more upon exposure to heat stress. Most proteins were reported to leave GroEL rapidly within 10-30 seconds. Three classes of substrate protein were distinguished: (i) proteins with a chaperonin independent folding pathway; (ii) proteins, more than 50% of total, with an intermediate chaperonin

dependence for which only a small fraction transits GroEL; and (iii) a set of highly chaperonin dependent proteins many of which dissociate slowly from GroEL and probably require sequestration of aggregation-sensitive intermediates within the GroEL cavity for successful folding.

Houry (Houry et al, 1999) extended the experiment mentioned above to determine GroE substrates *in vivo*. Pulse-chase labelling experiments were performed with live *E. coli* cells at 30°C in order to analyse the flux through GroEL of individual newly synthesised proteins. At different times of chase, cells were lysed on ice in the presence of EDTA, which prevented the ATP-dependent release of protein substrates from GroEL, and GroEL substrate complexes were isolated by immunoprecipitation with GroEL antibodies. Total soluble cytoplasmic proteins and GroEL bound proteins were separated on 2D gels. A number (~ 300) of newly synthesised polypeptides were reported to interact strongly with GroEL, some of which are structurally unstable and repeatedly rebind to GroEL. The identity of 52 (16%) of the polypeptides was determined by peptide-mass fingerprint analysis and MALDI-TOF mass spectroscopy; these included a number of essential components of the transcription/translation machinery, such as subunits of the RNA polymerase elongation factor Tu (EF-Tu) and several aminoacyl tRNA synthetases, as well as metabolic enzymes participating in a variety of important biosynthetic and catabolic pathways. From this data a structural motif was suggested for a 'typical' GroE substrate; it was suggested that a typical GroE substrate falls within the size range of 20-60 kDa and consists preferentially of multiple $\alpha\beta$ domains with hydrophobic β sheets. It was hypothesised that the non-native states of these topologically complex proteins are likely to expose extensive hydrophobic surfaces that could be recognised by GroEL, either during folding or upon misfolding. While these data suggest a possible architecture for GroE substrates, it is worth mentioning that 87% of all *E.coli* proteins are smaller than 50 kDa (Ewalt et al, 1997) and the $\alpha\beta$ topology is one of the most common folds found in proteins (~ 80% of all *E.coli* protein structures have at least one $\alpha\beta$ domain). From these figures, 69.6% of all *E.coli* proteins have a $\alpha\beta$ topology and are under 50 kDa! Therefore, while these experiments may have identified some GroE substrates, it may be a little premature to assign a typical architecture to GroE substrates.

1.2.5 Structure of GroES

GroES was isolated by Chandrasekhar *et al.* and characterized as a ring structure with approximately seven 10-kDa subunits that forms a complex with GroEL (Chandrasekhar et al, 1986). Crystallographic analysis of GroES has recently been completed [(Xu et al, 1997); Fig 1.2.1]. Each subunit consists of a β barrel region that forms most of the contacts around the ring, and a β-hairpin pointing slightly upwards and towards the centre of the ring. The β-hairpin region is loosely packed, with little intersubunit contact, and forms the roof of the dome-like structure of the oligomer. Earlier NMR work showed that a mobile domain in GroES became ordered upon binding to GroEL (Landry et al, 1993). This region (residues 17-32) is not seen on the X-ray structure of GroES, but is expected to form at least part of the binding contact. The visible ends of the mobile loop point downwards and radially outwards from the bottom of the β-barrel domain (dashed arrows, Fig 1.2.1b). EM reconstruction of the GroEL-GroES complex shows that the dome caps the opened apical domains of GroEL. In the complex, the mobile loop is expected to be in contact with GroEL, because its accessibility to trypsin is reduced in the complex and a synthetic peptide with the loop sequence binds to GroEL (Landry et al, 1993). Furthermore, all the GroES mutations originally isolated by their inability to support bacteriophage growth map to the mobile loop region of the sequence (Georgopoulos et al, 1973).

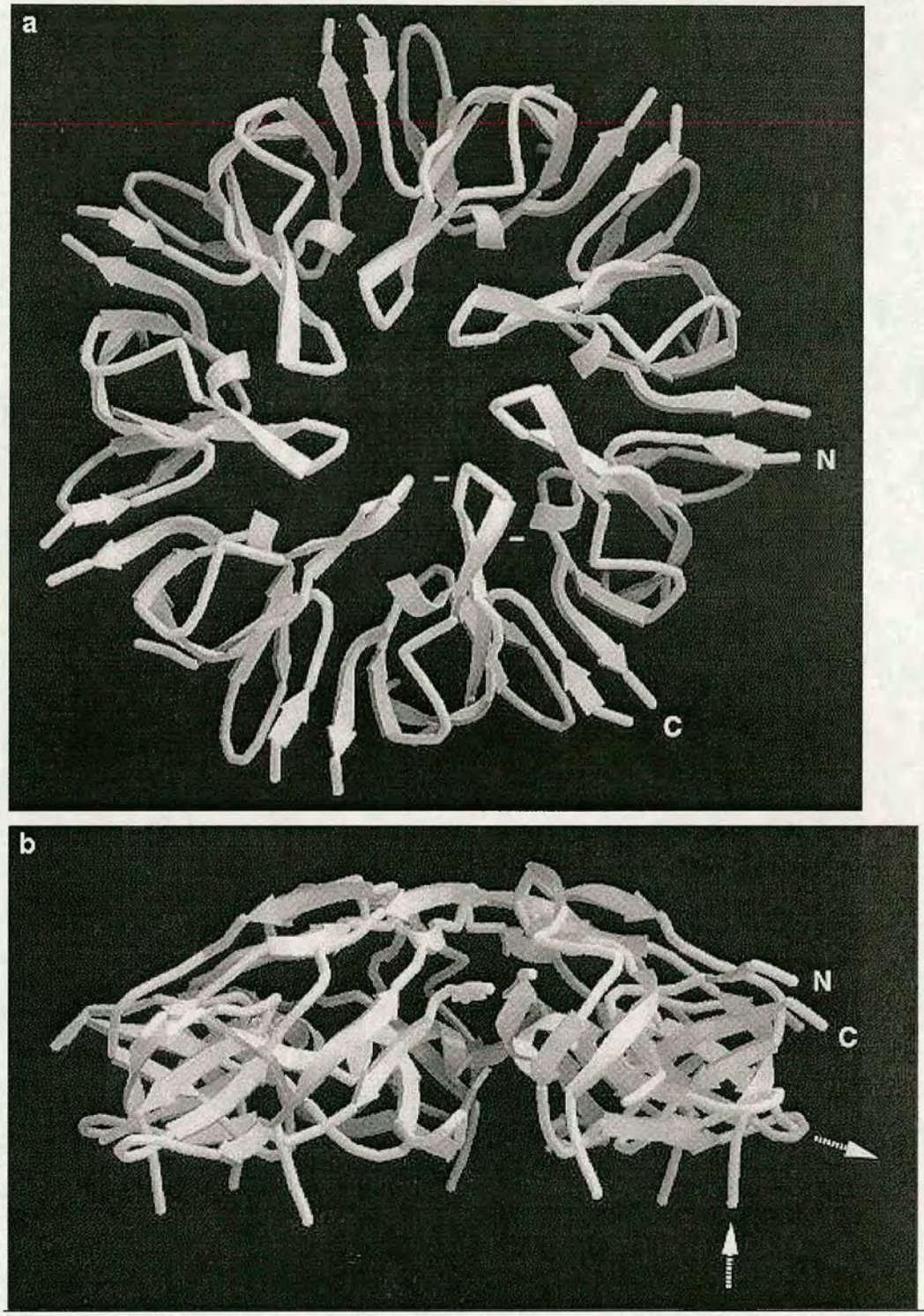


Fig. 1.2.1 Ribbon diagrams of the GroES structure viewed (a) from above and (b) from the side, showing the β-barrel structure of the subunit with a β-hairpin forming the roof of the dome-like heptamer. The N and C termini are labelled and point radially outward. The broken ends of the disordered mobile loop, which is not seen in this map, are indicated by dashed arrows in (b). Image from, The Chaperonins (Ellis, 1996).

1.2.6 Structure of GroEL

GroEL is a tetradecameric complex of identical 57 kDa subunits. Early electron microscopic analysis revealed its double toroidal structure, composed of two stacked heptameric rings with approximate dimensions of 14 nm in diameter and 16 nm in height (Hendrix, 1979; Hohn et al, 1979). This peculiar symmetry has been confirmed many times (Hutchinson et al, 1989; Langer et al, 1992; Zwickl et al, 1990), but, more recently, Saibil *et al.* performed a high resolution electron microscopic analysis of the Hsp60 from *Rhodobacter sphaeroides* (Saibil et al, 1993). This analysis enabled the clear visualisation of the two seven membered (7-mer) rings and supported the long held view that the Hsp60 monomer consisted of two major domains. The domains of each monomer are linked on the exterior of the cylinder, giving Hsp60 a cagelike structure (Saibil et al, 1993). The central cavity of the cylinder, intuitively the site of polypeptide binding, is 4 to 6 nm in diameter and can expand considerably upon binding of the GroES cofactor. Braig et al. accomplished the enormous task of solving the structure of the GroEL tetradecamer by X-ray crystallography (Braig et al, 1993) (Fig. 1.2.2). Each monomeric unit is composed of an apical, an intermediate, and an equatorial domain (Fig 1.2.3). The equatorial domain is the largest of the three and provides the majority of the subunit-subunit interactions and all of the interactions between the rings across the equatorial plane of the cylinder. The N and C terminal residues (about 30 residues) protrude into the cavity from the equatorial domain: they are disordered and are not crystographically resolved; however, the 'missing' 30 residues per subunit appear to form a central constriction to the channel as seen in cryo EM images (Chen et al, 1994). Mutation of Lys4-Glu (sequence beginning at Met1) completely blocks oligomer assembly (Horovitz et al, 1993), suggesting an important structural role for this apparently disordered region. Truncation of the carboxyl-terminus of GroEL by 16 amino acids resulted in a form of GroEL that hydrolysed ATP at a slower rate than wild type GroEL (McLennan et al, 1993). Mutagenesis later confirmed that this domain contains the ATP binding site (Fenton et al, 1994).

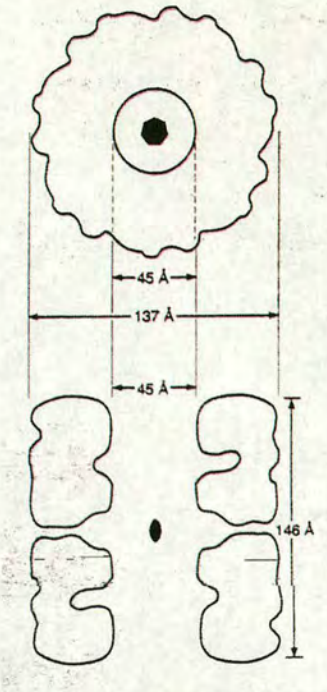


Fig. 1.2.2 Schematic model of the GroEL structure. Principal dimensions of the GroEL double toroid determined from the three dimensional crystal structure. The top panel shows a view down the central cavity, while the bottom panel represents a vertical cut through the cylinder. Image from Braig *et al.* (1994).

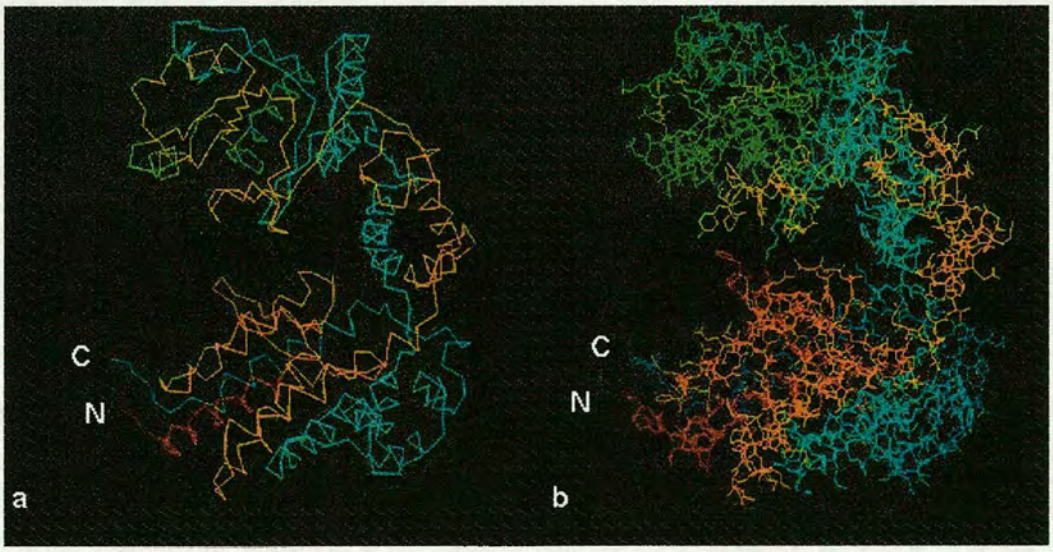


Fig 1.2.3 Atomic resolution structure of the GroEL subunit (Braig et al., 1994). (a) alpha-carbon and (b) full structure, colour coded in rainbow colours from N (red) to C (blue). In this view, the equatorial domain is the large mass containing both termini at the bottom of the subunit, connected by the antiparallel intermediate domain (cyan and yellow) to the large apical domain (green and cyan) at the top. Image from The Chaperonins (Ellis, 1996).

At the top of the equatorial domain, there is a well-defined junction with the intermediate domain, a small, antiparallel domain connecting the top and bottom large domains. It comprises a pair of crossed helices and a small β-strand region that contacts the neighbouring apical domain. This is followed by another junction leading to the large apical domain, which contains α and β structures and has poorly defined density in regions lining the channel, the site of residues involved in substrate and GroES binding, as identified by mutagenesis (Fenton et al, 1994). A conspicuous feature is that much of the apical domain surface is not in close contact with neighbouring domains, and that there is little barrier to its movement, both locally in some parts and in the overall orientation of the domain. It is this part of the structure that deviates most from 7-fold symmetry in the crystal. Very interesting potential hinge sites are found at the upper and lower domain junctions. These provide the possibility of large hinge movements allowing rigid body rotations (Gerstein et al, 1994) of the apical and equatorial domains. In these regions of exposed antiparallel chain, there are conserved glycine residues. Fig. 1.2.4 shows a schematic diagram representing the subunit domain arrangement in the oligomer, in which the potential hinge sites are marked with black dots.

An extensive mutational study (Fenton et al, 1994) has enabled the mapping of several important functional sites onto the atomic structure, including nucleotide, substrate and GroES binding sites. Results from other mutant studies (Baneyx et al, 1992; Yifrach et al, 1994) can now also be interpreted in terms of the 3D structure.

The helix bundle in the equatorial domain contains the ATP binding pocket, which is bordered by a highly conserved sequence motif, and lined by other stretches of highly conserved residues (Fenton et al, 1994). The structure with nucleotide bound has been solved (Xu et al, 1997), and confirms that nucleotide occupies this pocket. Mutations in Asp87 completely abolish ATPase activity, although ATP analogue binding is relatively unaffected. The ATP-binding pocket is adjacent to the lower hinge region, and mutations in the intermediate domain just beyond this hinge region, also abolish ATPase activity, as does mutation of residue 383, near the upper hinge region. These strategic locations of key residues around the hinge regions strongly suggest that hinge movements are involved in the hydrolysis mechanism.

The mutagenesis study defines a set of hydrophobic residues on the apical domain, lining the channel, that provide the binding surface for substrates, a region that is flexible in the GroEL structure. The binding surface is shown as a wavy profile on the apical domains in the diagram (Fig. 1.2.4a). Single amino acid substitutions of these residues abolish binding of the substrates ornithine transcarbamylase and dihydrofolate reductase. Density of the substrate malate

dehydrogenase (MDH) is found in exactly the same region, by cryo EM, held between the ends of the claw-like domains and protruding outwards (Chen et al, 1994). The ring of 7 binding sites is easily accessible from outside, being very near the surface of the cylinder, but the inward-facing orientation may protect GroEL from self-aggregation. Mutation of residue 152, near the lower hinge region, also has a strong effect on substrate binding. The ability of monomeric cpn60, missing 78 N-terminal residues, to bind substrates and partially promote their folding is consistent with the apical location of substrate binding sites (Taguchi et al, 1994).

Many of the same mutations to the apical channel-lining surface also interfere with GroES binding, suggesting that the binding sites for GroES and substrate overlap. In addition, the GroES binding surface appears to extend further over the top surface of the apical domain. Looking at the GroEL crystal structure, it would appear impossible to fit the GroES heptamer [diameter ~ 70 Å, Hunt *et al.*, in The Chaperonins (Ellis, 1996)] into the cavity to contact most of these binding sites. However, an explanation for this is found in the low-resolution structure of the GroEL-GroES complex from cryo EM (Fig 1.2.4b). A large rotation of at least the upper hinge region in the GroEL subunits is clearly evident, to give a reorientation of the apical domains, placing the substrate binding surfaces adjacent to the bases of the β-barrel domains of the GroES subunits (Fig 1.2.1b).

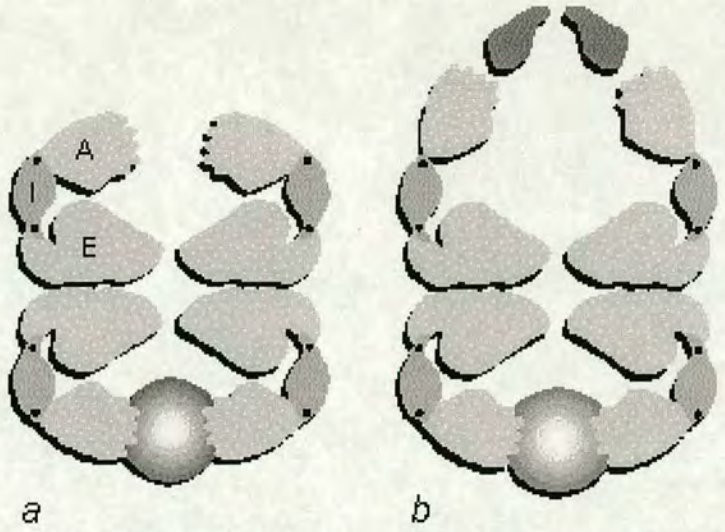


Fig. 1.2.4 Schematic diagram of the subunit arrangement in a hypothetical slice through the GroEL oligomer, showing the major functional sites. (a) GroEL, based on the crystal structure with bound substrate (shaded light to dark) as imaged by cryo-EM. A, apical domain; I, intermediate domain; E, equatorial domain. The apical domains form a ring of hand-like structures with the substrate binding sites on the fingers protruding into the central channel. Sites of potential hinge rotations are indicated by black dots. The notch in the equatorial domain represents the ATP-binding cleft. The interaction across the equatorial plane is shown as pairs of wavy surfaces. (b) The folding complex GroEL-MDH-GroES-ATP, rotating the subunit domains and adding the GroES (dark gray) and substrate (shaded) densities according to cryo EM observations. Image from The Chaperonins (Ellis, 1996).

1.2.7 The GroEL reaction cycle

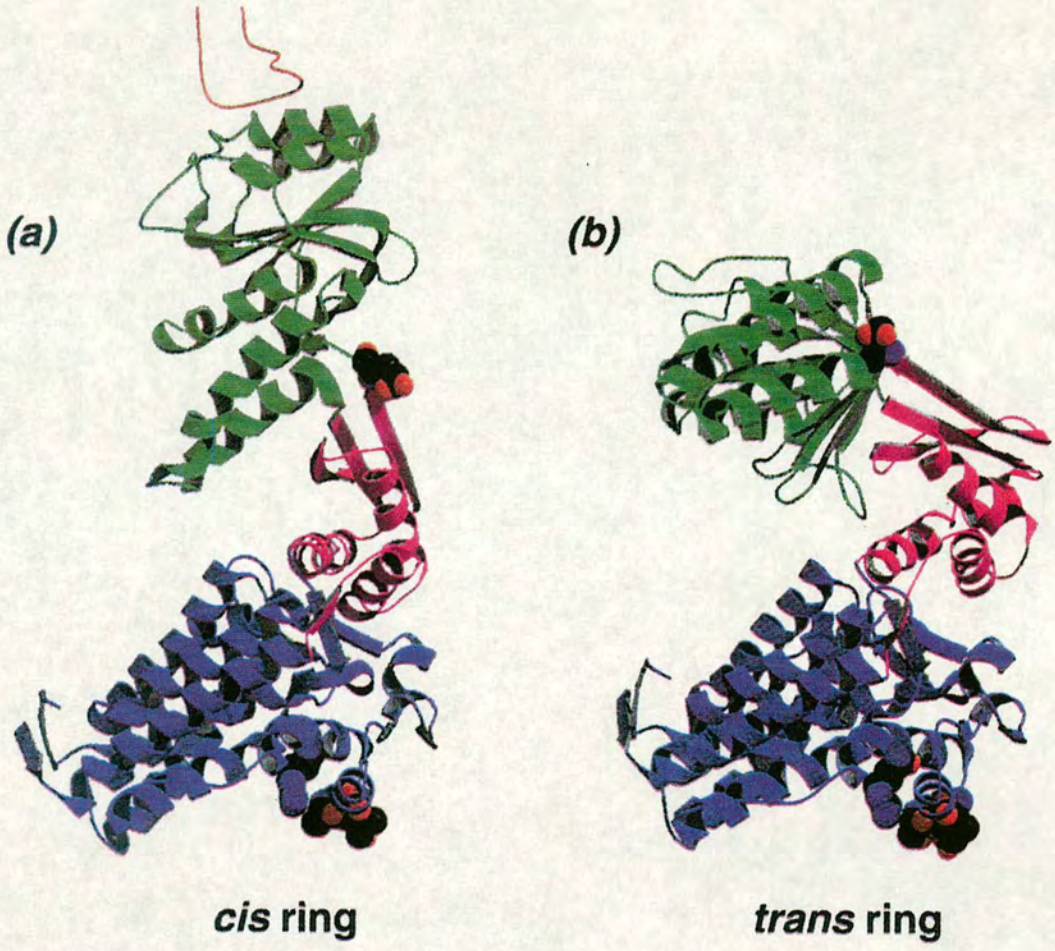
GroEL is thought to function as a double ring two-stroke complex (Lorimer, 1997). Both rings partake in a negative allosteric mechanism that synchronises the binding and release of GroES with the hydrolysis of ATP (Rye et al, 1997; Xu et al, 1997) to prevent misfolding and aggregation.

The double ring structure and the transmission of allosteric information between the rings of GroEL are thought to be key features for chaperonin activities (Rye et al, 1997; Xu et al, 1997). On the binding of GroES and ATP, the *cis* ring becomes expanded into a crown like structure (Roseman et al, 1996) doubling the volume of the GroEL cavity and hiding the hydrophobic binding surface (Xu et al, 1997). The apical domain moves 60° upwards and also rotates 90° (Fig 1.2.5), thus many of the GroEL hydrophobic residues that were involved in substrate binding are

now involved in the new GroEL inter-subunit interactions (Fig. 1.2.6). The thrusting apart of substrates via the ATP-driven rotation and translation of the apical and intermediate domains weakens further the binding of (Corrales et al, 1996) and forcibly unfolds, the protein substrates (Buckle et al, 1997; Shtilerman et al, 1999). An ATP-dependent conformational change in the opposite (*trans*) ring induces the release of GroES and, subsequently, of the folded protein substrates from the *cis* ring (Rye et al, 1997). GroEL undergoes ATP-induced conformational changes with a positive cooperativity between the subunits of each ring and with a negative cooperativity between the two rings (Yifrach et al, 1994). The binding of ATP to the *trans* ring, which primes the GroES for release, is possible only when the hydrolysis of ATP on the *cis* ring has occurred, i.e. the model of nested cooperativity (Yifrach et al, 1998; Yifrach et al, 2000). The reaction cycle of GroE is summarised in figure 1.2.7.

In support of the model described above, the single ring mutant ($GroEL_{SR1}$) of GroEL is functionally inactive *in vitro* (Weissman et al, 1996) and *in vivo* (Weber et al, 1998). This mutant contains four amino acid substitutions at the equatorial interface of the GroEL heptameric rings (Xu et al, 1997) that prevents the formation of double rings (Weissman et al, 1995). Effective folding of chemically denatured rhodanese occurs within the cavity of $GroEL_{SR1}$, but no release of GroES and of the folded enzyme is observed under physiological conditions (Weissman et al, 1996). The inability of $GroEL_{SR1}$ to release GroES is attributed to the absence of signal transmitted via the binding of ATP to the opposite ring (Weissman et al, 1996; Rye et al, 1997; Xu et al, 1997).

In contrast, a single-ring suffices for the mammalian mitochondrial homologue Hsp60 (Nielsen et al, 1998). Indeed, Hsp60 and the variants thereof created by fusing it with the equatorial domain of $GroEL_{SR1}$, function in *E. coli* as the sole GroEL (Nielsen et al, 1999). One possible explanation of this one-stroke engine could lie in the low affinity of Hsp60 for Hsp10, eliminating the requirement for ATP binding to a *trans* ring of the chaperonin in order to release the cochaperonin (e.g. in the case of GroEL/GroES). Furthermore, seven-membered single-ring minichaperones [i.e. fragments encompassing the apical domain of GroEL which are devoid of ATPase activity (Zahn et al, 1996)], have a slightly lower affinity for GroES compared with that of wild type GroEL and are more active than $GroEL_{SR1}$ (Chatellier et al, 2000b).



5

Fig. 1.2.5 A single subunit from the crystal structure of the asymmetric GroEL-GroES-$(ADP)_7$ chaperonin complex at 3.0Å resolution from (a) the *cis* ring and (b) the *trans* ring. The translation and rotation of the apical and intermediate domains serve both to increase the size of the GroEL-GroES cavity and to hide the substrate binding surfaces. The loop above image (a) represents the mobile loop of GroES. The R452E, E461A, S463A and V464A amino acids changes in $GroEL_{SR1}$are also highlighted. Image from Chatellier *et al.* (Chatellier et al, 2000a).

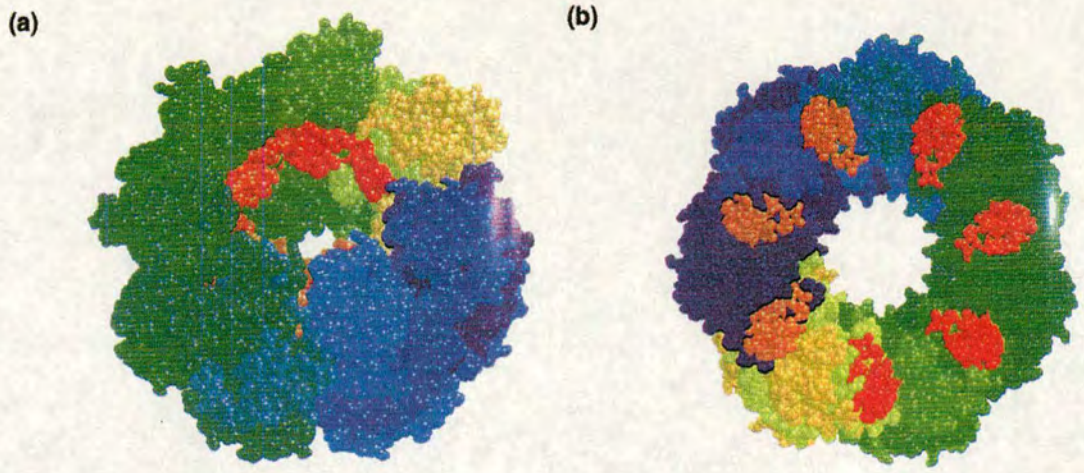


Fig. 1.2.6 One heptameric ring of the GroEL tetradecamer showing the seven polypeptide binding sites. (a) A large, continuous hydrophobic surface at the upper rim of the central cavity in the unliganded form. (b) Separated binding sites by the large scale movements of the apical domains upon the binding of nucleotide and GroES. GroEL subunits are coloured around the ring from blue to yellow, and the polypeptide binding sites are coloured from yellow to red. Image from Wang *et al.* (Wang et al, 2000)

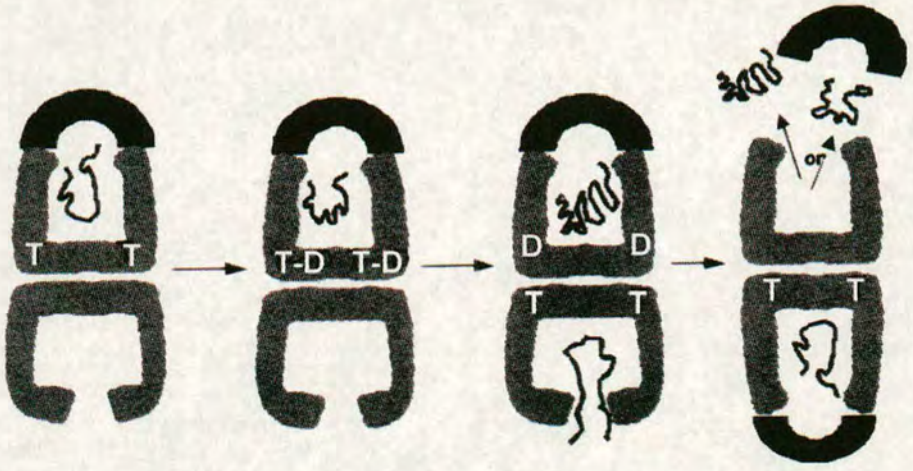


Fig. 1.2.7 Simplified representation of the GroEL folding cycle. T represents ATP and D represents ADP. GroES binds rapidly to the *cis* ring of GroEL after ATP binding. Following ATP hydrolysis to ADP in the *cis* ring, another substrate and ATP rapidly bind to the *trans* ring. Such binding results in the destabilisation and release of GroES and substrate from the *cis* ring. Those protein substrates that have not properly folded will rebind to the same or another GroEL ring. Image adapted from Ang *et al.* (Ang et al, 2000)

1.3 Mitochondrial Hsp60s

The eukaryotic mitochondrion is a membrane-bound organelle, and the site of several important metabolic pathways, the most notable being oxidative phosphorylation. Mitochondria possess their own genomes but most mitochondrial-associated proteins are actually encoded by nuclear genes, and subsequently imported into the organelle. Indeed some yeast cell can survive under certain conditions without mitochondrial DNA, but cells cannot survive without mitochondria, indicating the essential nature of some proteins that are imported into the mitochondria and their role within its confines.

A eukaryotic GroEL homologue was first seen as a heat inducible ribosome-associated protein in *Tetrahymena thermophila*. Antibodies raised against this 58 kDa protein showed that it was located in the mitochondria. Further analysis involving antibody cross-reactivity showed that the protein was indeed a Hsp60 homologue (McMullin et al, 1988). Hsp60 from *Saccharomyces cerevisiae* was cloned and sequenced, and shown to share 54% sequence identity with *E. coli* GroEL (Reading et al, 1989). Like *E. coli* GroEL, the protein demonstrates sevenfold rotational symmetry, exists as a homomeric tetradecamer, is induced by heat shock (although not as strongly as in *E. coli*) and is essential for cellular viability at all temperatures (Reading et al, 1989; Hallberg, 1990). Mitochondria are thought to be the remnants of a symbiotic relationship between a bacterium and the earliest eukaryotic cell; the finding of GroEL-related proteins solely in mitochondria (and chloroplasts in higher plants) goes some way to confirming this theory.

In chloroplasts, the cellular organelle in plants involved in photosynthesis, the GroEL homologue is the Rubisco subunit binding protein mentioned previously. It has been shown that this protein binds several imported polypeptides and its substrates are not just Rubisco subunits (Lubben et al, 1989).

While the range of *in vivo* substrates for the mitochondrial and chloroplast Hsp60s remain largely unknown, they are homologous to GroEL, having a requirement for both ATP hydrolysis and cofactor (HSP10) binding for the folding of polypeptides (Ranson et al, 1998).

1.4 The CCT chaperonins

Eukaryotic GroEL-homologues are found in cellular organelles. This has led to the search for molecular chaperones in the eukaryotic cytosol and the discovery of a second group of chaperonins. These group II chaperonins (Willison et al, 1994), found in archaebacteria and eukaryotes, are highly related to one another since they share up to 40% amino acid identity. They are weakly, but significantly, related in sequence to the group I chaperonins GroEL, Hsp60 and RSBP (Gupta, 1990; Ellis, 1990; Lewis et al, 1992). The archaebacterial chaperonins TF55 (Trent et al, 1991) and the thermosome (Phipps et al, 1993), are composed of only two-subunit species (Knapp et al, 1994) but the eukaryotic chaperonin CCT is composed of up to nine-subunit species (Rommelaere et al, 1993; Kubota et al, 1994). This situation is reminiscent of the 20S proteasome which is a multi-toroidal, multi-subunit complex involved in proteolysis; the archaebacterial proteasome is constructed from two subunits whereas the eukaryotic proteasome may have up to 14 different subunit types (Peters, 1994).

CCT is the only chaperonin identified so far in the eukaryotic cytosol and it has 16-18 subunits in a particle (Lewis et al, 1992). With the molecular cloning of all eight constitutively expressed subunits completed by Kubota's group (Kubota et al, 1995) the name was revised from TCP-1 complex (Lewis et al, 1992) to CCT, though it is also known as cytosolic chaperonin (Gao et al, 1992) or TriC (Frydman et al, 1992). It has been shown that CCT binds newly synthesised actin, tubulin and some other unidentified polypeptides *in vivo* (Sternlicht et al, 1993). CCT also redistributes in developing neurites of neuronal cells (Roobol et al, 1995) and binds a neurofilament peptide fragment *in vitro* (Roobol et al, 1993), suggesting that CCT may play a role in the organisation of the neuronal cytoskeleton. However, Lingappa (Lingappa et al, 1994) indicated that a TCP-1 like cytosolic chaperonin is involved in hepatitis B virus capsid assembly. These observations suggest that CCT may be involved in the folding and assembly of a wide range of cytosolic proteins.

As mentioned above, one of the striking differences between CCT and all the other chaperonins is the very heteromeric nature of the CCT particle. There are at least 7-9 major polypeptide constituents of CCT (Rommelaere et al, 1993; Kubota et al, 1994) and they are each independently encoded by highly diverged genes (Kubota et al, 1995). Because the other chaperonins only have one- or two-subunit species, this genetic complexity suggests that CCT may function as a more elaborate folding machinery that the other one or two-subunit-species chaperonins.

1.5 Previous work on GroE depletion in our laboratory

In an attempt to identify GroE substrates in *E. coli* a system was developed which would avoid the physiological consequences of shifting temperature (Kanemori et al, 1994), of using a mutant form of GroEL, or of burdening the cells with inactivated GroEL (Horwich et al, 1993). The system also allows the temperature to be retained at 37°C throughout and the synthesis of GroE to be completely shut off, allowing observation of the consequences of GroE depletion at optimum growth temperature. To do this a strain of *E. coli* was produced in which GroE synthesis was put under arabinose control (MGM100). Transcription of the *groELS* operon normally proceeds from a sigma-70 promoter and, after heat-shock, from a strong sigma-32 promoter (Kusukawa et al, 1988). This promoter region was replaced with the *araC* gene and *araBAD* promoter from pBAD-18 (Guzman et al, 1995), making production of GroE and continued viability of the construct dependant on added arabinose (McLennan et al, 1998).

The *araBAD* operon contains, moving upstream from the transcriptional start site, the *araI*, $araO_1$, and $araO_2$ control sites (Reeder et al, 1993). The *araI* site consists of two identical 17bp subsites, $araI_1$ and $araI_2$, which are direct repeats separated by 4bp, and are oriented such that $araI_2$, which overlaps the –35 region of the *araBAD* promoter, is downstream of $araI_1$. In the absence of AraC, RNA polymerase initiates transcription of the *araC* gene in the direction away from its upstream neighbour, *araBAD*. The *araBAD* operon is expressed at a low basal level (Fig 1.5.1 a).

When AraC is present, but neither L-arabinose nor CAP-cAMP (high glucose) (Kolb et al, 1993), AraC binds to $araO_1$, $araO_2$, and $araI_1$. $araO_1$ is the operator for the *araC* gene; its association with AraC blocks *araC* transcription so that this process is autoregulatory, although this requires high levels of AraC. The binding of AraC to $araI_1$ represses the expression of *araBAD* (negative control) (Fig. 1.5.1 b). A series of deletion mutations indicate that the presence of $araO_2$ is also required for the repression of *araBAD*. This remarkably large 211bp separation between $araO_2$ and $araI_1$ therefore suggests that the DNA between them is looped such that a dimeric molecule of AraC protein simultaneously binds to both $araO_2$ and $araI_1$. This is corroborated by the observation that the level of repression is greatly diminished by the insertion of 5bp (half a turn) of DNA between these two sites, thereby transferring $araO_2$ to the opposite face of the DNA relative to $araI_1$ in the putative loop. Yet the insertion of 11bp (one turn) of DNA has no such affect (Lobel et al, 1990).

When L-arabinose is present, it allosterically induces the AraC subunits bound to $araO_2$ to instead bind $araI_2$. This activates RNA polymerase to transcribe the *araBAD* genes (positive control). When the cAMP level is high (low glucose) (Kolb et al, 1993), CAP-cAMP binds to a site between $araO_1$ and $araI_1$ (although in most other CAP-binding operons, CAP-cAMP binds adjacent to RNA polymerase), where it functions both to help break the loop between $araO_2$ and $araI_1$ and to increase the affinity of AraC for $araI_2$. *araC* remains repressed by AraC-L-arabinose at $araO_1$ (Schleif, 1992) (Fig 1.5.1 c).

Experiments on MGM100 showed that an arabinose concentration of 0.2% was found to give normal growth both on plates and in liquid, and close to normal levels of GroE (Fig. 1.5.2). After arabinose removal, and the addition of 0.2% glucose, GroE levels halve with each generation (Fig. 1.5.2). Figure 1.5.3 shows that when arabinose is removed, and glucose added, growth of MGM100 remains exponential for about 1.5 hours (GroE levels will be reduced to 2-3% of wild type level by this time). Growth continues at a decreasing rate for another hour (at this time, GroE levels will be reduced to ~ 0.25% of wild type level); the cells then abruptly lyse.

Lysis indicated the development of a defect in the cell envelope. To determine if the cellular membranes remained intact the GroE depleted cells were grown in osmotically protected medium. In the presence of 10 mM Mg^{2+} lysis did not occur, also the normally rod-shaped cells were converted to spherical protoplasts instead of lysing completely. This result was characteristic of a defect in synthesis of the peptidoglycan cell wall, responsible for the rigid rod-shape of *coli* cells (Heijenoort, 1996). It was therefore suspected that an enzyme or enzymes concerned with the synthesis or maintenance of the peptidoglycan layer required GroE for synthesis or continued integrity.

Damage to the peptidoglycan layer could result either from failure of peptidoglycan synthesis or enhanced peptidoglycan breakdown (controlled breakage and resealing of peptidoglycan bonds is necessary to allow cell growth). In order to distinguish between these two possibilities, the rate of incorporation of ^{3}H-diaminopimelic acid (DAP), a specific precursor, into peptidoglycan (Höltje et al, 1975) was measured. It was expected that uptake of DAP would either be reduced (if synthesis of peptidoglycan was affected) or constant (if degradation of peptidoglycan was affected). It was therefore surprising that DAP uptake was actually greatly increased in GroE depleted cells (Table 1.5.1), suggesting that the cells may be DAP starved.

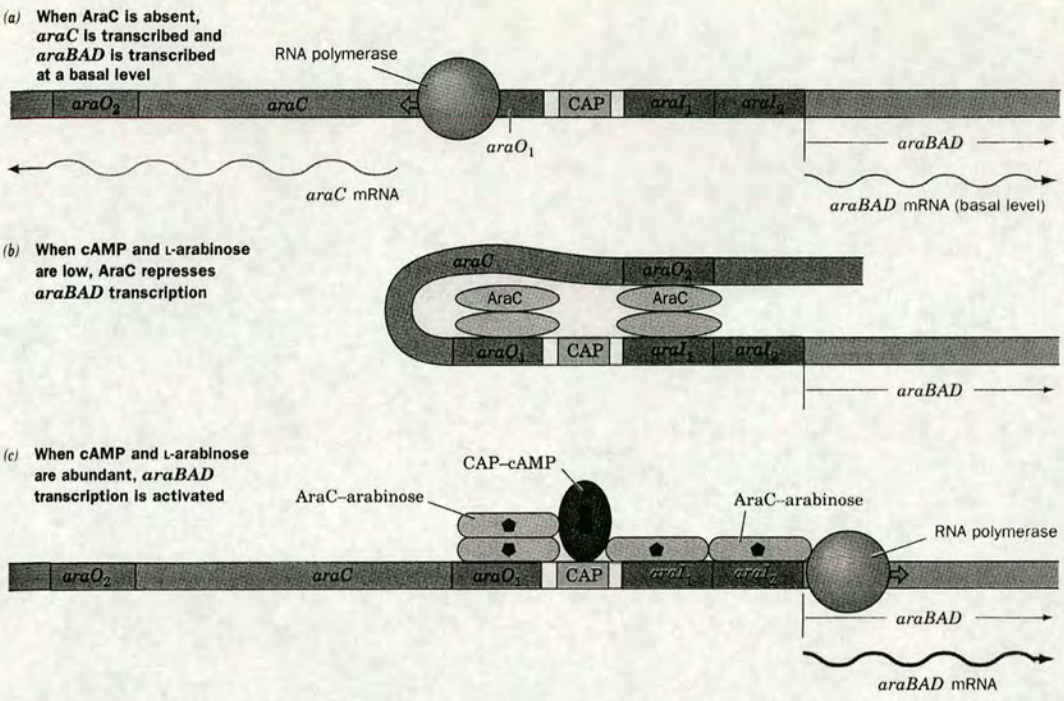


Fig. 1.5.1 The proposed mechanism for *araBAD* regulation: (a) in the absence of AraC, RNA polymerase initiates the transcription of *araC*. *AraBAD* is also expressed at a low basal level. (b) When AraC is present, but not L-arabinose or cAMP, AraC binds to $araO_1$ and links together $araO_2$ and $araI_1$ to form a DNA loop, thereby repressing both *araC* and *araBAD*. (c) When AraC and L-arabinose are both present and cAMP is abundant, CAP-cAMP stimulates the AraC-arabinose comples to release $araO_2$ and instead bind $araI_2$, thereby activating *araBAD* transcription. *AraC* remains repressed. (Image taken from Biochemistry, 2nd ed. Voet and Voet. John Wiley & Sons, Inc. pub.)

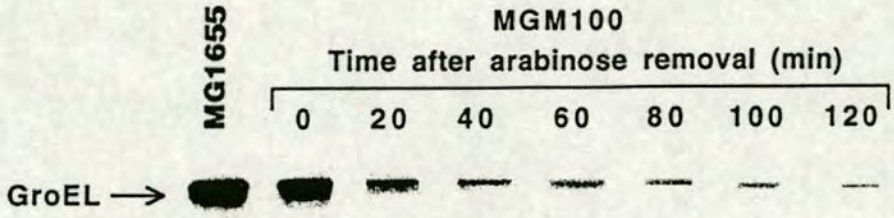


Fig 1.5.2 GroEL depletion after arabinose removal. MGM100 grown in the presence of 0.2% arabinose produces slightly less GroEL than does MG1655 (~ 85%). Image from McLennan and Masters (McLennan et al, 1998)

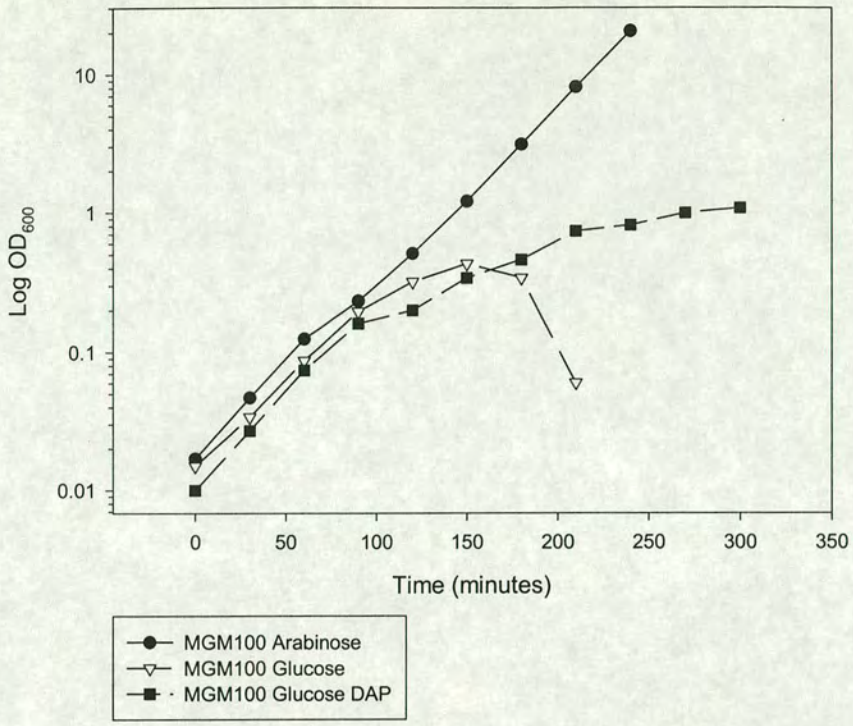


Fig. 1.5.3 Growth curves of MGM100 growing in LB + arabinose, glucose or glucose + DAP.

Time after arabinose removal (min.)	Relative counts per minute per cell
0	1.0
85	1.0
140	2.1
170	3.1
200	6.5
230	7.4
255	7.6

Table 1.5.1 Incorporation of ^{3}H-DAP into peptidoglycan by GroE depleted cells. Data from N. McLennan (pers. comm.).

In confirmation of this it was found that the addition of DAP to the growth medium completely prevents cell lysis (Fig. 1.5.3), although exponential growth does not persist for longer than in DAP free broth. Western analysis showed that the addition of DAP did not increase or stabilise GroEL levels. Furthermore, although addition of DAP to solid medium did not allow colony formation in the presence of glucose, confirming that GroE is essential for processes other than cell wall formation, when DAP is present, colony formation requires less arabinose than normally (Table 1.5.2). This suggested that at least part of the cellular complement of GroE is dedicated to ensuring that DAP is available and that GroE is likely to have a role in maintaining the levels of an enzyme or enzymes involved in DAP synthesis.

Further confirmation that cell wall synthesis is adversely affected in GroE depleted cells came from an experiment in which the enzyme which adds DAP to the growing peptidoglycan precursor is overproduced. It was reasoned that, if this enzyme (not implicated in DAP-deficiency induced lysis because it uses rather than synthesizes DAP) were expressed from a high copy number plasmid extra enzyme would be present before GroE depletion, which might partly compensate for reduced levels of DAP. This proved to be so. It was shown that lysis is delayed in depleted cells with extra copies of the *murE* gene (MurE is the enzyme which adds DAP to UDP-N-acetylmuramyl-L-alanine-D-glutamic acid). Providing extra copies of *murD*, *murF* or *mraY*, other enzymes of murine monomer synthesis, which do not use DAP as a substrate (Heijenoort, 1996) did not delay lysis.

The same approach was used to discover at what step in DAP synthesis there might be a deficiency of enzyme resulting from GroE depletion. DAP is synthesized in a 6-step pathway from aspartic semialdehyde (Cohen et al, 1987). Deficiency of any of the enzymes DapA-E has been reported to lead to cell lysis (Patte, 1996). It was hypothesized that if a critical enzyme were to be synthesized in excess before GroE depletion, lysis might be delayed. At the time of the experiments the position of *dapC* had not been defined, but PCR was used to separately clone *dapA*, *B*, *D*, and *E* into pBR325. The presence of the *dapA* plasmid was shown to delay lysis after GroE depletion by several hours (Fig. 1.5.4) while the other *dap* plasmids had no delaying affect. This suggested that GroE might influence levels of DapA. Western blotting was used to look at the fate of DapA in GroE depleted cells. Cultures were sampled at 160 minutes after arabinose removal, just before lysis occurs. At this time DapA levels were reported to be reduced to 16% of normal.

Arabinose concentration (%)	Number of colonies	
	+ DAP	-DAP
0.2	404	400
0.1	383	394
0.08	455	151
0.07	474	78
0.06	396	5
0.05	85	1
0.04	0	0

Table 1.5.2 The presence of DAP reduces the cells' requirement for GroE. Data from N. McLennan (pers. comm.).

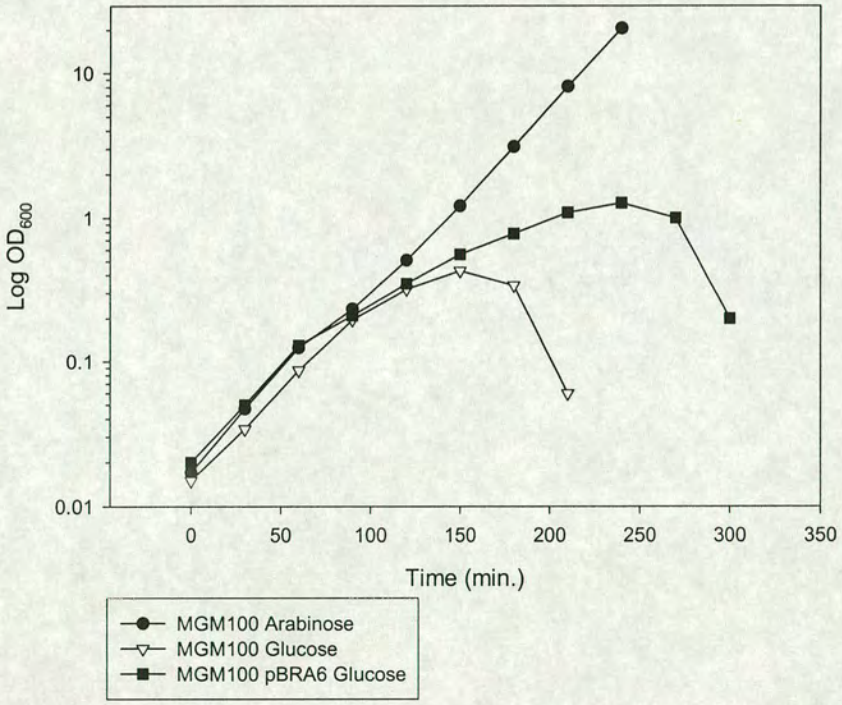


Fig 1.5.4 Cloned *dapA* (pBRA6) delays the onset of GroE depletion induced lysis. Data from N. McLennan (pers. comm.).

1.6 The lysine biosynthetic pathway

The work described in this thesis is mainly a continued investigation into the effects of GroE depletion on DapA; therefore DapA and the lysine biosynthetic pathway will be discussed here in further detail.

Work's discovery in 1950 of the unusual amino acid *meso*-diaminopimelate in acid hydrolysates of *Corynebacterium diphtheriae* (Work, 1950), and subsequently in *Mycobacterium tuberculosis* (Work, 1951) led to the systematic study of its biosynthesis and the description of the bacterial lysine biosynthetic pathway (Gilvarg, 1960). Not only was diaminopimelate found to be the immediate precursor of L-lysine, but it was subsequently demonstrated to be a component of the bacterial cell wall (Cummins et al, 1956): specifically, the peptidoglycan portion. In *Escherichia coli*, and other bacteria, not only L-lysine, but also L-threonine, L-methionine, and L-isoleucine are derived from L-aspartate (Fig. 1.6.1) The common portions of the biosynthetic pathway include the enzymes aspartokinase and aspartate semialdehyde dehydrogenase, which generate aspartate semialdehyde. The reduction of aspartate semialdehyde by homoserine dehydrogenase yields homoserine, the precursor to threonine, methionine and isoleucine.

Bacteria have evolved three slightly different strategies for the biosynthesis of diaminopimelate, and thus lysine, from aspartate (Scapin et al, 1998). After generation of aspartate semialdehyde, it is condensed with pyruvate by the action of the *dapA*-encoded synthase to generate dishydropicolinate. This compound is reduced by the *dapB*-encoded reductase to yield tetrahydropicolinate. The acyclic form of tetrahydropicolinate, L-2-amino-6-keto-pimelate, can be directly converted to diaminopimelate by the reduced nicotinamide adenine denucleotide (NADPH) and ammonia-dependent reductive amination of the 6-keto group, a reaction catalysed by diaminopimelate dehydrogenase (Misono et al, 1976). Diaminopimelate dehydrogenase has a very limited occurrence, having only been identified in *Corynebacterium glutamicum* (Misono et al, 1986), *Bacillus sphaericus* (White, 1983) and a *Pseudomonas* and *Brevibacterium* species (Misono et al, 1986). As an alternative to the direct reductive amination reaction, two alternative pathways exist in bacteria, which involve the *N*-acylation of L-2-amin-6-ketopimelate to prevent cyclisation. By far, the most prevalent pathway involves *N*-succinylation of the α-amino group by succinyl CoA, catalysed by the tetrahydropicolinate/succinylCoA *N*-succinyltransferase (Fig. 1.6.2). Once the acyclic form is stabilized by succinylation, a pyridoxal phosphate (PLP) dependent transaminase generates a second amino acid

center of the L configuration, using L-glutamate as the amino group donor. The product of the transaminase, *N*-succinyl-L-L-diaminopimelate is desuccinylated by the *dapE*-encoded desuccinylase to generate succinate and L-L-diaminopimelate. The epimerisation of this intermediate form to D-L-diaminopimelate is catalysed by the *dapF*-encoded diaminopimelate epimerase, one of the rare examples of a non-PLP-dependent amino acid epimerase. While many amino acid biosynthetic enzymes are clustered into multigenic operons, including the enzymes of isoleucine biosynthesis, the genes encoding the lysine biosynthetic enzymes are not clustered, and in fact are scattered over both the *E. coli* and *Haemophius influenzae* circular chromosomes (Bukhari et al, 1971).

The alternate pathway involving *N*-acetylated intermediates instead of *N*-succinylated intermediates proceeds through an identical series of steps, and is, like the dehydrogenase pathway, extremely limited in its bacterial distribution. The deacylation of *N*-acetyldiaminopimelate has only been described in *Bacillus* species, with the acetylase pathway of *B. megaterium* being the best characterized system (Sundharadas et al, 1967).

In contrast to bacterial lysine biosynthesis, yeast synthesise L-lysine via the α-ketopadipate pathway, in which acetyl CoA and α-ketoglutarate are condensed to form homocitric acid (Strassman et al, 1964). Isomerisation generates homoisocitric acid (Strassman et al, 1966), which is oxidized to oxaloglutaric acid and decarboxylated to yield α-ketoadipate (Strassman et al, 1965). α-Ketoadipate is transaminated to α-aminoadipate (Broquist et al, 1961), reduced to α-aminoadipate semialdehyde (Sagisaka et al, 1961), and condensed with L-glutamate to generate saccharopine (Jones et al, 1966) named for the yeast *Saccharomyces cerevisiae*, from which the compound was first identified. Oxidation and hydrolysis of the imine yields L-lysine and α-ketoglutarate (Saunders et al, 1966).

In the following sections the enzymes involved in lysine biosynthesis in *Escherichia coli* shall be described in greater detail.

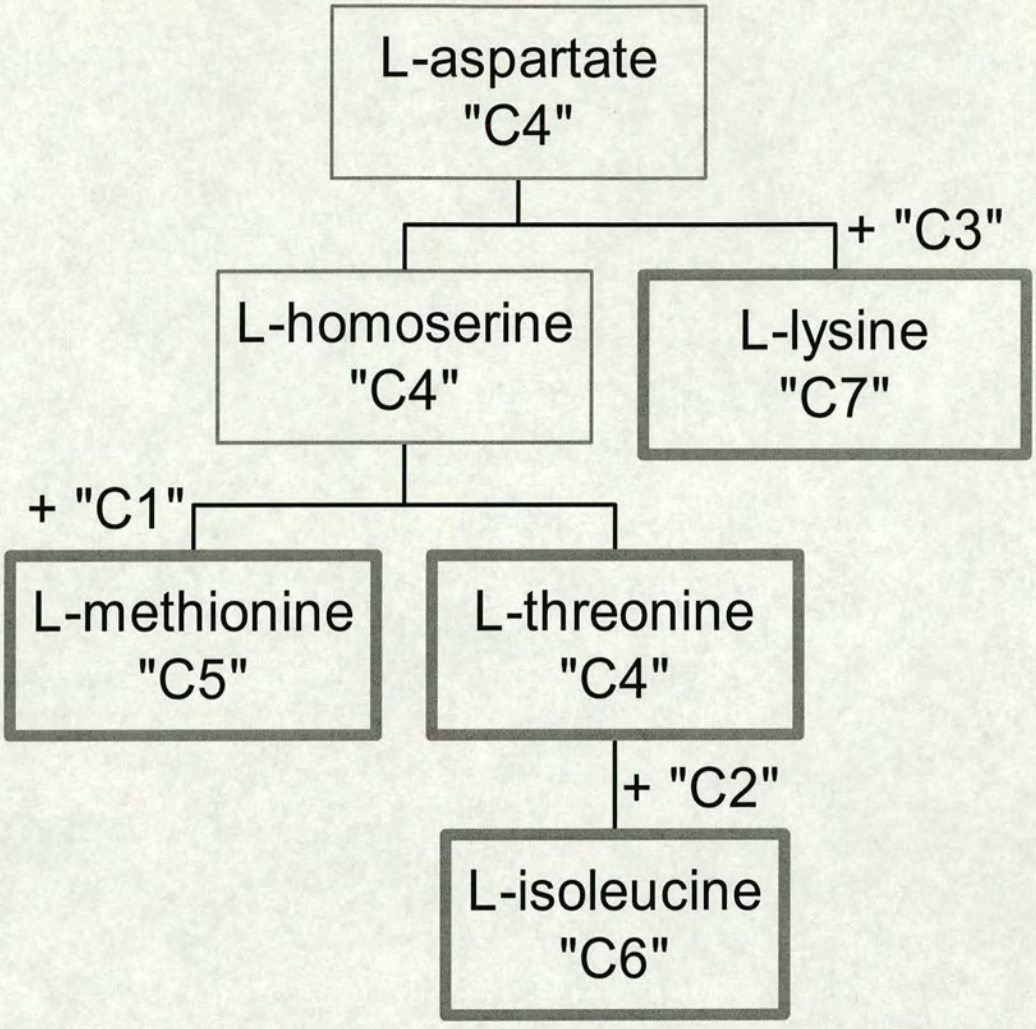


Fig. 1.6.1 Bacterial amino acid biosynthesis using aspartate as a precursor. Heavily outlined boxes are essential amino acids in mammals.

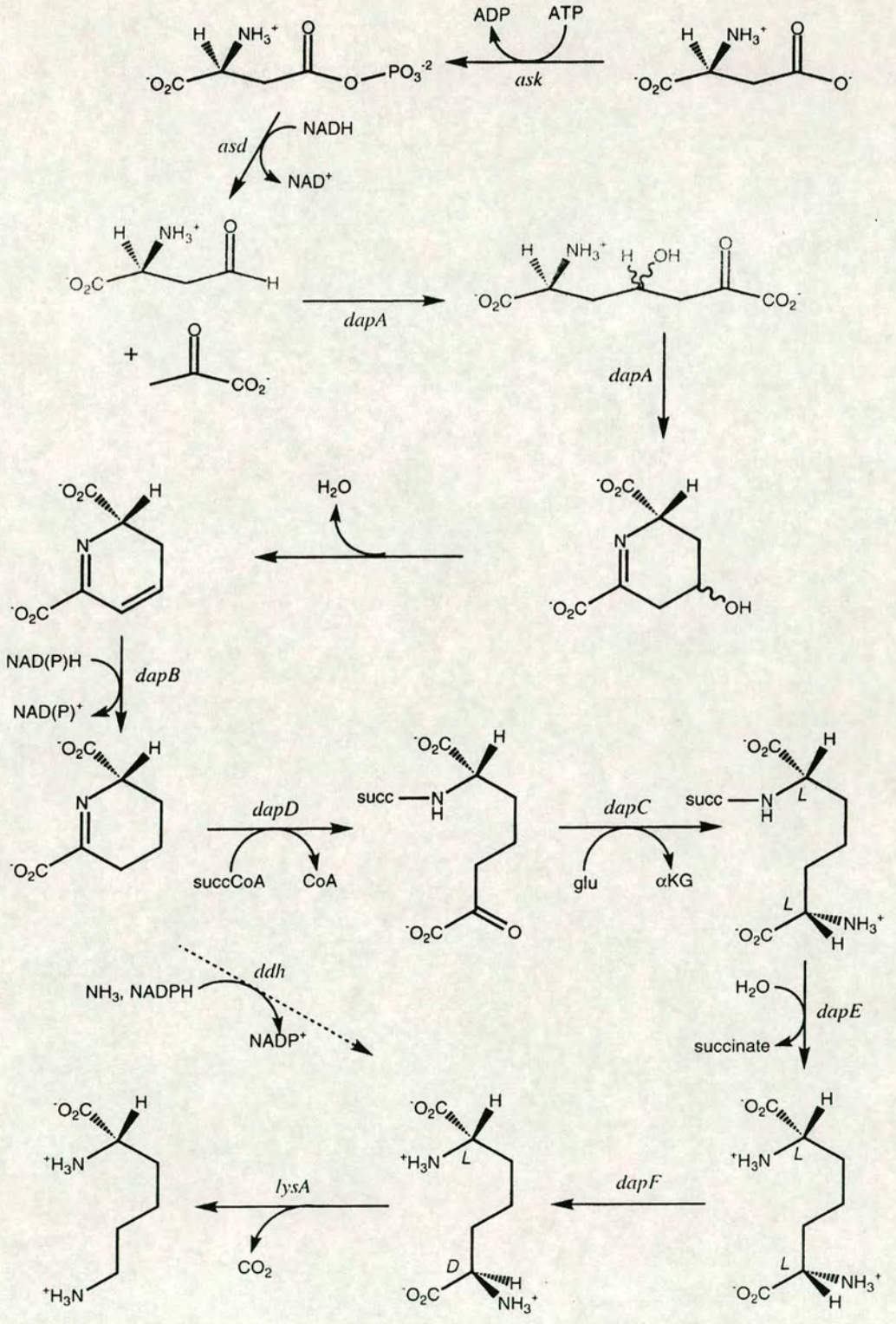


Fig. 1.6.2 Bacterial L-lysine biosynthetic pathway. Image from Scapin and Blanchard (Scapin et al, 1998)

1.6.1 Aspartokinase and aspartate semialdehyde dehydrogenase

In all species of bacteria, aspartokinase and aspartate semialdehyde dehydrogenase (EC 2.7.2.4 and 1.2.1.11) act sequentially to generate the early, common intermediates in L-threonine, L-isoleucine, L-methionine and L-lysine biosynthesis. Aspartokinase has been the subject of intense genetic and biochemical interest due to the presence of three isozymes in *E. coli*, which are end-product inhibited by the three amino acid products: L-threonine, L-methionine and L-lysine respectively (Patte, 1996). The most abundant isozyme is the *thrA*-encoded aspartokinase I, a bifunctional aspartokinase-homoserine dehydrogenase (Stadtman et al, 1961). This 820 amino acid *E. coli* protein is composed of an amino terminal aspartokinase domain and a carboxy terminal homoserine dehydrogenase domain (Truffa-Bachi et al, 1968). Threonine exhibits competitive inhibition versus aspartate and noncompetitive inhibition versus homoserine, but the inhibition curves are both sigmoidal suggesting cooperative binding to the homotetramer (Cohen, 1985). Threonine and isoleucine also regulate the expression of the *thrA*-encoded aspartokinase. The *E. coli metL*-encoded aspartokinase II is also a bifunctional aspartokinase-homoserine dehydrogenase, which in contrast to aspartokinase I, is not inhibited by methionine (Patte et al, 1967). The enzyme exists as a homodimer of 809 amino acid long monomers, and expression of the gene is repressed by methionine. Finally, the *lysC*-encoded aspartokinase III is a monofunctional aspartokinase, which is noncompetitively and cooperatively inhibited by lysine and other amino acids (Patte et al, 1965; Stadtman et al, 1961). The *E. coli* aspartokinase III is a homodimer of 449 amino acid long monomers that are homologous to the amino terminal domains of the aspartokinase I and II.

The *asd*-encoded aspartate semialdehyde dehydrogenase (EC 1.2.1.11) catalyses the reversible NADPH-dependent reduction of aspartyl phosphate to generate aspartate semialdehyde. The *E. coli* gene has been sequenced (Haziza et al, 1982), and encodes a 367 amino acid monomer that exists as a homodimer in solution (Biellmann et al, 1980). Homology is observed between the Asd protein from *E. coli* and other sequences from different species (Chen et al, 1993).

1.6.2 Dihydropicolinate Synthase

The *dapA*-encoded dihydropicolinate synthase (DHDPS) (EC 4.2.1.52), located at min 53 on the chromosome (Bukhari et al, 1971) catalyses the first unique step in lysine biosynthesis: the aldol condensation between pyruvate and aspartate semialdehyde. The activity was first described in extracts of *E. coli* in 1965 (Yugari et al, 1965), and the enzyme was purified 5000-fold to apparent homogeneity from *E. coli* in 1970 (Shedlarsky et al, 1970). Since it catalyses the committed step in L-lysine biosynthesis it is perhaps not surprising that all synthases studied to date are feedback inhibited by L-lysine. The extent of inhibition has been used to group the synthases into three classes: gram positive bacterial synthases are only modestly inhibited by high concentrations of lysine; plant synthases are strongly inhibited at lysine concentrations less than 50 μM; while the *E. coli* and *B. sphaericus* enzymes are weakly inhibited by 0.2-1.0 mM concentrations of L-lysine (Patte, 1996). Early kinetic studies suggested that L-lysine was a competitive inhibitor versus aspartate semialdehyde, but more recent studies with the wheat enzyme suggest that lysine is a bona fide allosteric inhibitor (Shaver et al, 1996).

The native *E.coli* enzyme was reported to have a molecular weight of 134 kDa, and is now known to be a homotetramer of 31,372 Da monomers (Richaud et al, 1986). These early studies identified a lysine residue as being involved in the initial binding of pyruvate. A ping-pong mechanism has been proposed for DapA function; pyruvate binds to the enzyme, forming a Schiff base with the ε-amino group of Lys-161, and is followed by the binding of aspartate semialdehyde (Laber et al, 1992).

The three-dimensional structure of the *E. coli* dihydropicolinate synthase has been solved using X-ray crystallography at 2.5 Å resolution (Mirwaldt et al, 1995). The asymmetric unit is composed of two monomers, and the tetramer is generated using crystallographic symmetry operations. The tetramer has a large solvent-filled cavity at its center (Fig. 1.6.3). Each monomer is composed of two domains: the amino terminal domain assumes an eight-stranded parallel α/β (TIM) barrel ranging from residues 1-224, while the carboxy-terminal domain is composed predominantly of α-helices ranging from residues 224-292 (Fig. 1.6.4). The catalytic site, as definied by the position of Lys-161, is at the carboxyl-terminal end of the barrel, with the ε-amino group of Lys-161 almost exactly centered in the barrel. This site is the position of all catalytic sites in such parallel α/β barrel enzymes (Farber et al, 1990), and is commonly used to stabilize developing negative charges on intermediates, or in the transition state, for enzyme-catalysed reactions. In a

continuation of these studies Farber and Petsko (Farber et al, 1990) successfully soaked a number of substrates and substrate analogues into the apoenzyme, and determined the structure of many of these complexes (Blickling et al, 1997). The pyruval complex of the enzyme showed the expected formation of the covalent Schiff base with Lys-161.

Of substantial interest was the determination of the structure of the L-lysine complex of DHDPS. Two molecules of L-lysine were shown to bind at the monomer-monomer interface, with interactions between each lysine molecule and residues of both monomers potentially explaining the observed cooperative binding of L-lysine to the tetramer (Blickling et al, 1997). The α-carboxyl group of bound L-lysine interacts predominantly with the phenolic oxygen of Tyr-106, and the α-amino group makes extensive interactions with Asn-80 and Glu-84. The ε-amino group makes interactions with His-53 and His-56, and all of these residues are observed to move from their position in the apoenzyme upon lysine binding. Sequence alignments (Mirwaldt et al, 1995) of the derived protein sequences from plants and bacteria show that while Asn-80 and Tyr-106 are conserved in all synthases, Glu-84 is only present in synthases that exhibit strong lysine inhibition, and is replaced by a threonine residue in gram-positive bacteria that are weakly inhibited by lysine. In support of the critical role of Glu-84 in L-lysine binding to the synthase, is the report that mutations of this residue in the maize synthase result in the synthesis of a synthase that has lost sensitivity to L-lysine inhibition (Shaver et al, 1996). This finding suggests that mutant synthases that are not feedback inhibited by lysine may be useful to generate transgenic plants with increased levels of free lysine for animal feedstock, allowing for this required, and limiting, amino acid to be more easily acquired.

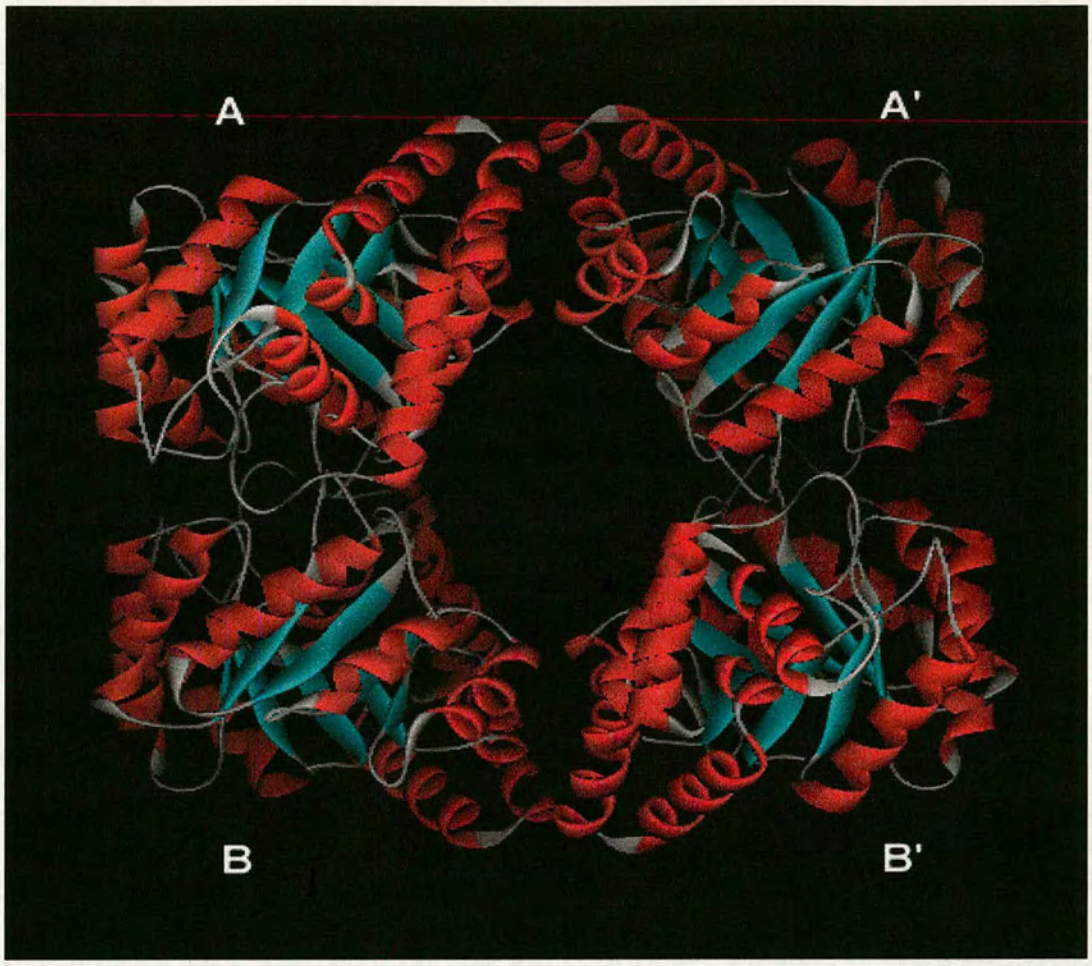


Fig. 1.6.3 Three-dimensional structure of *E. coli* dihydropicolinate synthase tetramer. (Coordinate filename 1DHP.pdb)

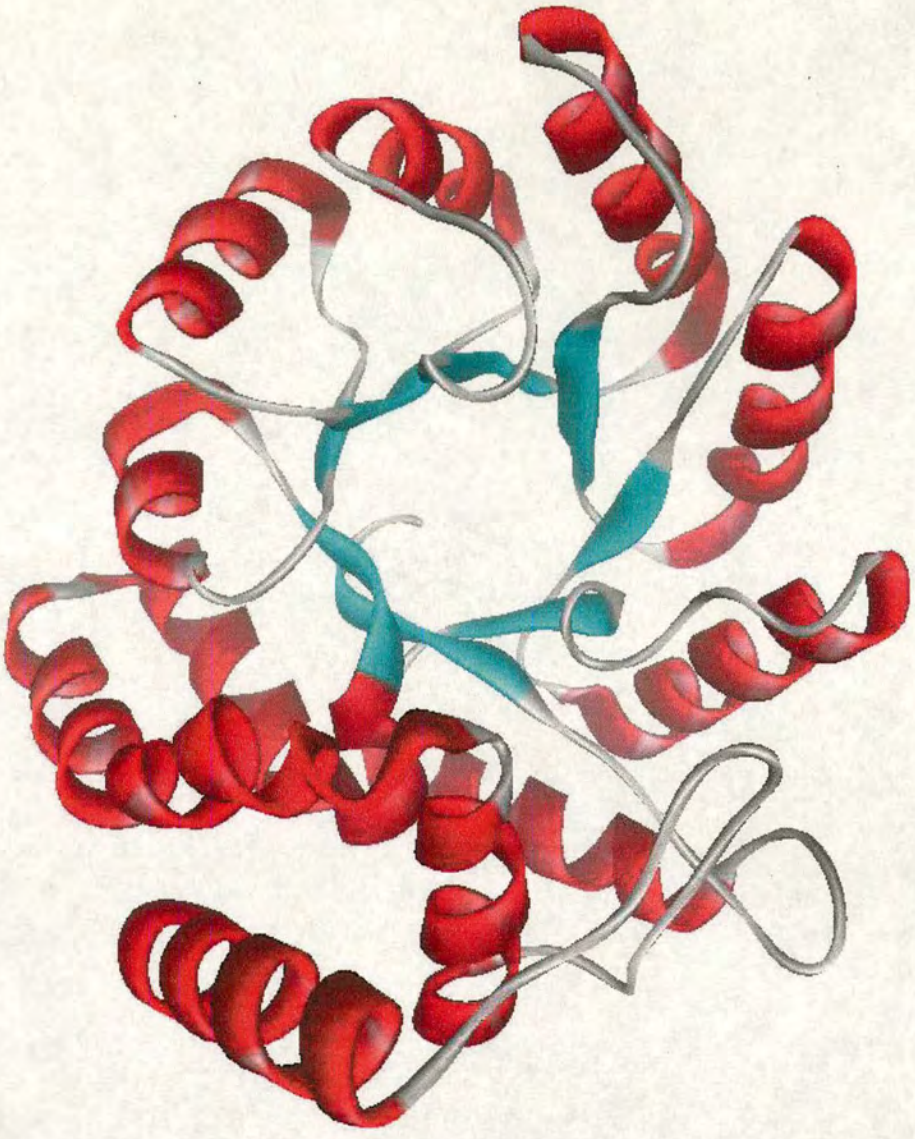


Fig. 1.6.4 Three-dimensional structure of *E. coli* dihydropicolinate synthase monomer. (Coordinate filename 1DHP.pdb)

1.6.3 Dihydropicolinate Reductase

The *dapB*-encoded dihydropicolinate reductase (EC 1.3.1.26), located at min 0.5 on the chromosome (Bukhari et al, 1971), catalyses the thermodynamically favourable NAD(P)H-dependent reduction of dihydropicolininc acid (DHDP) to tetrahydropicolinic acid (THDP), the second step in the lysine biosynthetic pathway. The reductase was initially identified in *E. coli* mutants that required diaminopimelic

acid and lysine in their growth medium, and were subsequently show to lack the reductase (Farkas et al, 1965). The enzyme was subsequently purified 1900-fold (Tamir et al, 1974) to homogeneity. The enzyme exhibits a broad pH optimum around 7, and the enzyme's native molecular weight was found to be 110 kDa by gel filtration chromatography and sedimentation velocity measurements. The enzyme was shown to have a high affinity for its substrates, with reported K_m values of 9 and 10 μM for dihydropicolinate and NADPH, respectively.

The first *dapB* gene sequence was reported in 1984 for the *E. coli* enzyme (Bouvier et al, 1984). The DNA sequence predicted a 273 residue polypeptide, with a subunit mass of 28,798 Da. To date, eight additional sequences have been reported (Scapin et al, 1998). They all encode reductases of between 246 and 286 residues. All share several regions of homology, in particular the motif (^{8}V/I)(A/G)(V/I)XGXXGXXG^{18}, located at the extreme N-terminus of the protein, and the motif ^{157}E(L/A)HHXXKXDAPSGTA^{171} (numbers are for the *E. coli* sequence). The first motif is found in a variety of dinucleotide-dependent dehydrogenases, and in the three dimensional structure of the *E. coli* enzyme is part of the nucleotide binding domain (Scapin et al, 1995). The second region has been suggested as a potential substrate-binding region on the basis of the conserved positively charged residues and its location in the three-dimensional structure of the *E. coli* enzyme (Pavelka et al, 1997).

1.6.4 Tetrahydropicolinate N-Succinyltransferase

The *dapD*-encoded tetrahydropicolinate *N*-succinyltransferase (EC 2.3.1.117) is localised at 3.5 min clockwise from *tonA* (Richaud et al, 1984). The gene has been cloned and sequenced, and encodes an open reading frame of 274 amino acids (monomer molecular weight 30,040 Da) (Richaud et al, 1984).

The native *E. coli* enzyme has been purified and extensively characterised (Simms et al, 1984). The purification to homogeneity required a 1900-fold purification from wild type *E. coli*, and the subunit molecular weight by SDS_PAGE analysis was about 31 kDa. The purified enzyme was shown to catalyse the reversible transfer of the succinyl group from succinyl CoA to tetrahydropicolinate, although the forward, physiologically important reaction was 380 times faster than the reverse reaction at the pH optimum of 8.2 (Simms et al, 1984).

The three dimensional structure of the *N*-succinyltransferase has been reported at 2.2 Å (Beaman et al, 1997). Primary sequence studies had previously

noted that the transferase was homologous to a family of enzymes that bore a signature motif, termed the 'hexapeptide repeat', or leucine patch (Vaara, 1992). The principal feature of this motif is the presence of multiply repeated sequences, (LIV)-(GAED)-X-X-(STAV)-X, which has been structurally characterised in two previous cases: the *E. coli* UDP-*N*-acetlyglucosamine acetlytransferase (Raetz et al, 1995) and the *Methanosarcina thermophila* carbonic anhydrase (Kisker et al, 1996). The homologous hexapeptide repeat sequences of both these enzymes form an unusual left-handed β-helix. Both enzymes exist as trimers, as does the *dapB*-encoded tetrahydropicolinate *N*-succinyltransferase. The active site was proposed to be at each monomer-monomer interface, since this is where both pCMBS (a mercurial) and cobalt (II), inhibitors of the reaction, were bound. More recent studies of the enzyme crystallised in the presence of 2-aminopimelate [identified as an alternate substrate (Berges et al, 1986a)] and CoA have confirmed the presence of these two molecules at the monomer-monomer interface.

1.6.5 N-Succinyl-L-L-Diaminopimelate Aminotransferase

Almost nothing is known about the aminotransferase (DapATase) (EC 2.6.1.17) that catalyses the transfer of the amino group from L-glutamate to *N*-succinyl-L-2-amino-6-ketopimelate to generate ketoglutarate and *N*-succinyl-L-L-diaminopimelate. Furthermore, no sequence of any *dapC* gene has been reported from any source, although several reports have incorrectly identified the *dapD* gene as the *dapC* gene (Richaud et al, 1984). The enzyme was first partially purified from *E. coli* (Peterkofsky et al, 1961), and could be purified approximately 110-fold (Peterkofsky, 1962). More recently, the enzyme has been purified 1500-fold to homogeneity (Cox et al, 1996). The enzyme has a subunit mass of 39,896 Da, and exists as a homodimer. The enzyme is absolutely specific for L-glutamate as the amino donor, but will use a number of *N*-acyl ketopimelates.

A recent report suggests that the *argD*-encoded *N*-acetylornithine aminotransferase (NAcOATase) exhibits both NAcOATase and DapATase activity, with similar specificity constants for *N*-acetylornithine and *N*-succinyl-L-L-diaminopimelate, suggesting that it can function in both lysine and arginine biosynthesis (Ledwidge et al, 1999). This finding may explain why numerous investigations have failed to identify genetically the bacterial *dapC* locus.

1.6.6 N-Succinyl-L-L-Diaminopimelate Desuccinylase

Along with the preceding enzyme in the lysine biosynthetic pathway, the *dapE*-encoded *N*-succinyl-L-L-diaminopimelate desuccinylase (EC 3.5.1.18), localised at 53 min, is poorly characterised. The substrate of the enzyme was first identified in 1959 (Gilvarg, 1959), but the enzyme was not purified until about 10 years ago (Gelb et al, 1990). Enzyme levels do not appear to be regulated by lysine levels in the media, and the enzyme was purified to homogeneity from *E. coli* K12 grown in rich medium. The enzyme was purified 7100-fold, and was found to be a mixture of dimeric and tetrameric 40 kDa monomers. The enzyme requires a divalent metal ion for activity, with cobalt being more effective than zinc on the basis of a 2.2-fold higher maximum velocity in the presence of cobalt (Lin et al, 1988).

The amino acid sequences of five bacterial desuccinylases have been reported, and alignment of the sequences from *E. coli* (375 amino acids), *H. influenzae* (377 amino acids), *C. glutamicum* (369 amino acids), *H. jannaschii* (410 amino acids), and *M. leprae* (337 amino acids) reveals a highly homologous consensus sequence in the first 100 residues (Boyen et al, 1992). This consensus pattern has also been identified in several other metal-dependent enzymes that catalyse the hydrolysis of amide linkages (Meinnel et al, 1992). These include the *E. coli argE*-encoded acetylornthinase, the *Pseudomonas cpg2*-encoded carboxypeptidase G2, the *Lactobacillus pepV*-encoded carnosinase, the bacterial *pepT*-encoded tripeptidase, and both the yeast *yscS*-encoded carboxypeptidase and the mammalian aminoacylase-1 (Boyen et al, 1992). This finding suggests that the conserved amino acid sequence may be structurally important for metal ligation and catalysis. The three dimensional structure of the *Pseudomonas cpg2*-encoded carboxypeptidase G2 has been determined to 2.5 Å (Rowsell et al, 1997). The structure revealed the presence of two zinc atoms at the presumptive active site that share a bridging water/hydroxide ligand, and also demonstrated that the conserved glutamate, aspartate and histidine residues were involved in metal ligation.

1.6.7 L-L-Diaminopimelate Epimerase

The *dapF*-encoded diaminopimelate epimerase (EC 5.1.1.7), located at min 85 (Richaud et al, 1987), catalyses the interconversion of L,L-DAP, the product of the *dapE*-encoded desuccinylase reaction, and D,L-DAP, the substrate for the *lysA*-encoded decarboxylase. The enzyme was initially identified in crude extracts of *E.*

coli in 1957 by Work and co-workers using paper chromatographic identification of the formation of D,L-DAP from L,L-DAP (Antia et al, 1957) and only purified to homogeneity from *E. coli* in 1984 (Wiseman et al, 1948). In this latter study, formation of D,L-DAP from L,L-DAP was followed spectrophotometrically by coupling the product to the diaminopimelate dehydrogenase, which is specific for the D,L-isomer. The homogenous enzyme was purified 6000-fold in 3% yield and shown to behave as a monomer on gel permeation chromatography. The subunit molecular weight determined by SDS-PAGE was 34 kDa, as subsequently confirmed by the cloning and sequencing of the gene from *E. coli* (Richaud et al, 1988). Although earlier studies had yielded ambiguous data on the presence of pyridoxal phosphate as a cofactor in the epimerisation reaction, the homogenous enzyme exhibited no visible absorbance indicative of the presence of PLP, neither was the reaction stimulated by exogenously added PLP, nor inhibited by hydrazine or hydroxylamine, reagents known to inhibit PLP-dependent enzymes (Wiseman et al, 1948). Studies with the active site directed aziridine derivative, 2-(4-amino-4-carboxybutyl)-2-aziridine-carboxylic acid, demonstrated the irreversible and covalent inactivation of the epimerase. Tryptic digestion followed by peptide separation and sequencing demonstrated that Cys-73 was alkylated by this compound (Higgins et al, 1989). This residue is conserved in the aligned sequences of four bacterial *dapF*-encoded epimerises (Scapin et al, 1998).

Few effective inhibitors have been designed and tested against the purified epimerase. It is clear that the *dapF*-encoded epimerase is not an essential gene in *E. coli*, since its disruption has no obvious lethal consequences to the organism (Richaud et al, 1987). While *dapF* mutants accumulate L,L-DAP intracellularly, it has been proposed that other enzymes may act to epimerise the L,L-DAP to generate sufficient D,L-DAP for incorporation into the peptidoglycan component of the cell wall. As described above, the aziridine derivative is a potent, active site directed irreversible inhibitor of the epimerase, although its antibacterial properties have not been explored.

1.6.8 D-L-Diaminopimelate Decarboxylase

The activity of the *lysA*-encoded D,L-diaminopimelate decarboxylase (EC 4.1.1.20) was first demonstrated in crude extracts of *E. coli* in 1952 (Dewey et al, 1952). The enzyme has been purified from a number of bacteria (White et al, 1965) and plants, and catalyses the only PLP dependent decarboxylation of an α-amino

acid of D configuration (Scapin et al, 1998). In addition to this unique substrate stereospecificity, it has been shown that both the *B. sphaericus* and wheat germ decarboxylases catalyse the decraboxylation reaction with inversion of configuration (Asada et al, 1981; Kelland et al, 1985). This is in contrast to all other stereochemical outcomes for PLP-dependent decarboxylases.

Essentially no *bona fide* mechanistic studies have appeared for diaminopimelate decarboxylase, in spite of renewed interest in the potential antibacterial properties of inhibitors of this enzyme.

In the last twenty years, a number of sequences of bacterial *lysA* genes have appeared, including those from *E. coli* (Stragier et al, 1983a), *C. glutamicum* (Yeh et al, 1988), *P. aeruginosa* (Martin et al, 1988), *B. subtilis* (Yamamoto et al, 1989) and *M. tuberculosis* (Andersen et al, 1993). In addition, the *lysA* genes of *H. influenzae* and *M. jannaschii* have been identified from genome sequencing projects by homology to those sequences reported above (Fleischmann et al, 1995; Bult et al, 1996; Bult et al, 1996). Alignment of these bacterial decarboxylase sequences reveals three conserved lysine residues, one of which is likely to be the lysine that binds PLP cofactor. The bacterial diaminopimelate decarboxylase sequences are also homologous to other basic amino acid decarboxylases, including both bacterial and mammalian ornithine decarboxylases and arginine decarboxylase (Scapin et al, 1998). An extensive alignment of PLP-dependent enzymes has been used to separate all such enzymes into four types, based on primary sequence homology (Alexander et al, 1994). Diaminopimelate decarboxylase is a member of the third type, which includes eukaryotic ornithine decarboxylases, plant and bacterial arginine decarboxylases, and bacterial alanine racemases (Grishin et al, 1994). Both sequence alignment and secondary structural predictions have been used to suggest that the type 3 PLP-containing enzymes assume an eight stranded α/β barrel. The recent determination of the high-resolution three-dimensional structure of *B. stearothermophilus* alanine racemase (Shaw et al, 1997) provides support for this prediction. The racemase consists of an amino-terminal α/β barrel domain composed of residues 1-240, and a carboxyl-terminal composed of β strands. The lysine residue involved in Schiff base formation with the cofactor is located at the carboxyl terminus of the first of the eight central beta strands that make up the barrel.

The chemistry catalysed by PLP-dependent enzymes has made them the subject of numerous mechanism-based inhibitor studies. Based on similar studies of ornithine decarboxylase, fluoromethyl diaminopimelates have been synthesised and shown to be inhibitors of the epimerase (Osterman et al, 1995). The β,γ-unsaturated *meso*-diaminopimelate has been synthesised and shown to be a modest inhibitor of

the decarboxylase (Girodeau et al, 1986). More recently, inhibitors in which the carboxylic moiety undergoing decarboxylation has been replaced by the corresponding phosphonate have been prepared and shown to inhibit the decarboxylase, as well as the dehydrogenase and epimerase (Song et al, 1994). Unfortunately, these compounds are both poor inhibitors of the enzyme and exhibit negligible antibacterial activity against *E. coli* or *Salmonella*. However, tripeptides containing the phosphonate analogue exhibit some antibacterial activity, reminiscent of successful antibacterial studies in which L-2-aminopimelate was incorporated into tripeptides, presumably to enhance transport of these molecules across the cell wall and membrane (Berges et al, 1986b). The determination of the three-dimensional structure of the decarboxylase may allow some of these studies to be resuscitated.

1.6.9 Regulation of lysine biosynthesis

It has been clearly established, either directly or by the use of fusions with *lacZ* that the expression of several genes in the lysine biosynthetic pathway depends on the intracellular pool of lysine; this is true for *lysC* (Stadtman et al, 1961), *asd* (Boy et al, 1972), *dapB* (Bouvier et al, 1984), *dapD* (Richaud et al, 1984) and *lysA* (Boy et al, 1979a). In contrast, expression of *dapA* (Butour et al, 1974), *dapE* and *dapF* (Patte, 1996) appears constitutive. Results are yet to be reported for *dapC*. However, even though regulation by lysine has been demonstrated, the molecular mechanism that correlates variations of lysine concentration with variations of gene expression remains to be established and differs quite certainly from one gene to another.

The expression of LysC specific activity varies widely (over 100-fold between repression values and lysine-limited chemostat values) with the lysine intracellular pool. While a number of mutational studies have been performed on the *lysC* gene (Boy et al, 1979b) and the promoter region has been sequenced (Cassan et al, 1983), it is still not understood how lysine plays a role in its regulation.

Expression of the *asd* gene varies in response to the concentration of lysine, threonine and methionine (Boy et al, 1972), but lysine is a much more important effector (50-fold variation) than the other two (only 2-fold derepression) and therefore is included in the lysine regulon. Operon fusions with *lacZ* have indicated that regulation occurs at the transcriptional level (Patte, 1996). Observations of the sequence (Haziza et al, 1982) has led to the conclusion that a typical multivalent attenuation mechanism is not involved [such a mechanism exists for the

Streptococcus asd gene (Cardineau et al, 1987)]. The molecular mechanism of regulation therefore remains unknown. It has been shown that another metabolite, glucose 6-phosphate, modifies *asd* expression (Boy et al, 1979c), but again, no molecular mechanism has been proposed.

Variations in the lysine pool also greatly affect *dapB* expression [20-fold derepression (Boy et al, 1978)]. Mutations in the *relA* gene (Patte et al, 1980) and in the *lysS* gene (Boy et al, 1976) indicate that ppGpp on the one hand and lysyl-tRNA on the other must play a role in this regulation. However, examination of the nucleotide sequence upstream of the *dapB* gene did not allow the proposal of a molecular mechanism. Some sequence similarities with the *asd* promoter sequence have been suggested (Bouvier et al, 1984). Binding of an unidentified cytoplasmic protein to the DNA promoter region has been observed in gel retardation experiments (Patte, 1996).

Similarly, examination of the coding sequence upstream of the *dapD* gene gave no indication of the molecular mechanism by which lysine modifies the expression of this gene (Richaud et al, 1984).

The only example in which a clear mechanism has been established is the regulation of the *lysA* gene. Expression of this gene depends on the intracellular concentration of both diaminopimelate, which acts as an inducer, and lysine, which acts as a corepressor. This double regulation reflects the special situation of diaminopimelate decarboxylase, as a catabolic enzyme for diaminopimelate and an anabolic enzyme for lysine (Boy et al, 1979a). Moreover, ppGpp must play a role in *lysA* expression, which is modified in *relA* mutants (Patte et al, 1980). Mutants with a Lys phenotype were obtained outside the *lysA* structural gene. They had lesions in the *lysR* gene, specifying a positive regulatory protein for *lysA* expression (Stragier et al, 1983b). The *lysR* gene is located next to *lysA* and transcribed divergently (65 bp lie between the two sites of transcription initiation). Genetic evidence has shown that LysR (without diaminopimelate) binds to an intergenic sequence, allowing activation of the *lysA* promoter (Stragier et al, 1983a) and repression of the *lysR* promoter (autoregulation). It has been proposed that lysine blocks LysR binding, but this point remains to be firmly established. The *lysR* sequence has been determined and found to encode a protein of 34 kDa (Stragier et al, 1983c). A characteristic helix-turn-helix sequence for DNA binding has been demonstrated.

The respective roles of feedback inhibition and repression have also been studied, selecting, for example, mutants resistant to growth inhibition by analogues of lysine (*S*-aminoethyl cysteine). The LysC step is a crucial point of regulation for lysine biosynthesis. The feedback inhibition is 100%. Only desensitised *lysC* mutants

could be isolated (with the exception of permeation mutants), and accumulation of lysine in the growth medium results from such mutations (Boy et al, 1972).

The biological importance of feedback inhibition on the DapA enzyme has been questioned, in view of the high K_i of the ligands. However, it must play a role *in vivo*, in parallel with some channelling, as desensitised *lysC* mutants overproduce mainly lysine, and not threonine.

Finally, *in vivo* determination of the limiting steps in the lysine pathway was made possible by the use of mutants and by gene amplification with multicopy plasmids carrying one of the genes of the pathway: the LysC activity is the first point of limitation of the flux, followed by the DapA step (Daucé-Lereverend et al, 1982)

1.6.10 Summary

The widespread occurrence of the bacterial biosynthetic pathway described in the preceeding sections has suggested that inhibition of the enzymes in this pathway may have antibacterial properties (Girodeau et al, 1986). The absence of the enzymes in this pathway in mammals and the dual biosynthetic importance of the pathway, in both protein and cell wall biosynthesis, bolster these arguments. Recent insertional inactivation of the *ask* gene has demonstrated that the biosynthesis of diaminopimelate is an essential function in *Mycobacterium smegmatis*, and that organisms lacking the *ask* gene product suffer 'DAP-less' death (Pavelka et al, 1996). The demonstrated clinical effectiveness of inhibitors of enzymes involved in cell wall biosynthesis, particularly peptidoglycan synthesis, suggests that a more thorough enzymological and structural characterisation of these unique enzymes may yield new classes of inhibitors with therapeutic efficacy. The recent appearance of multiply antibiotic-resistant bacterial pathogens increases the urgency to find new and effective antibacterial compounds, with broad spectrum antibacterial activity.

1.7 The project

Previous attempts have been made to determine the *in vivo* substrates of GroE in *Escherichia coli* using a number of different approaches (Horwich et al, 1993; Kanemori et al, 1994; Houry et al, 1999). In a further attempt to identify GroE substrates, McLennan and Masters constructed a strain of *E. coli*, in which expression of the chaperonin GroE could be regulated by the presence or absence of exogenous arabinose (McLennan et al, 1998). It was found that when *E. coli* are GroE depleted they lyse, due to diaminopimelate starvation. It was further discovered that diaminopimelate was not being synthesised, as the protein DapA is not functional in GroE depleted cells. The work described in chapter 3 of this thesis is a continuation of the study of the effects of GroE depletion on DapA in *E. coli*. Work in chapter 4 describes the analysis of another protein, which appears to be affected by GroE depletion, the lac repressor.

During the initial investigations into DapA, it was found that expression of the *dapA* gene increased during GroE depletion. This was an unusual result, as *dapA* was not believed to be regulated at the level of expression (Bouvier et al, 1984). Further investigation revealed that, contrary to earlier reports, *dapA* is in fact regulated at the level of expression, and this is discussed in detail in chapter 5.

Chapter II

Materials and Methods

2.1 Bacterial strains, phage and plasmids

2.1.1 Growth Media and buffers

Bacterial strains used in this study are listed in Table 2.1.1. Bacteria were either maintained on L-Broth plates stored at 4°C, or for longer-term storage, in frozen storage buffer [50% bacterial buffer + 50% glycerol (v:v)] at -70°C.

Yeast strains used in this study are listed in Table 2.1.2. Yeast was maintained on YPD or DOB plates stored at 4°C.

Bacteriophages used in this study are listed in Table 2.1.3. Phage lysates of P1 and λ were stored at 4°C as broth suspensions to which a few drops of chloroform had been added to prevent microbial growth.

Plasmids used and constructed in the course of this study are listed in Table 2.1.4.

Growth media and commonly used buffers are listed in Tables 2.1.5 and 2.1.6 respectively. L-Broth and L-agar were used routinely for all bacterial manipulations, except where otherwise stated. For work with phage λ the medium was supplemented with 10 mM $MgSO_4$ and 0.2% maltose to maximise expression of the λ receptor protein; for phage P1, 2.5 mM $CaCl_2$ was added. VB minimal agar, supplemented with appropriate carbon sources, vitamins and amino acids (concentrations are defined later in the text), was used for selection of nutritional markers. For work with MGM100 the media were supplemented with 0.2% arabinose or 0.2% glucose and DAP was added at a final concentration of 100μg ml^{-1}.

2.1.2 Growth of Bacteria

Bacteria were routinely grown as liquid cultures at 37°C (or 30°C for temperature-sensitive strains). Usually, fresh overnight cultures that had been inoculated from a single colony were diluted the following day and grown as required.

Table 2.1.1 Bacterial strains used in this study

Strain	Genotype	Source/Reference
AG31	F^-, *groES*131^{ts}	C. Georgopolous
BL21(DE3)	E. coli B strain. F^-, *ompT*, *hsdS*$_B$ (r_B-m_B-), λDE3	(Studier et al, 1986)
CAG18439	F^-, *lam*$^-$, *rph*-1, *lacZ*118(Oc), *lacI*3042::Tn*10*(Tet^R)	(Singer et al, 1989)
CAG18442	F^-, *lam*$^-$, *rph*-1, *thrA*34::*Tn*10(Tet^R)	(Singer et al, 1989)
DH5α	ϕ80*lacZ*ΔM15, Δ(*lacZYA-argF*)U169, *recA*, *endA*, *hsdR*17 (r_K- m_K+), *supE*44, λ-, *thi*-1, *gyrA*96, *relA*1	Laboratory stock
ET505	F^-, *lam*$^-$, *lysA*::*Tn*10(Tet^R), IN (*rrnD-rrnE*)1	Ishiguro, E.E.
INV110	F^-, (*tra*Δ36, *proAB*, *lacI*^{q}, *lacZ*ΔM15) *rpsL* (str^R), *thr*, *leu*, *endA*, *thi*-1, *lacY*, *galK*, *galT*, *ara*, *tonA*, *tsx*, *dam*, *dcm*, *supE*44, Δ(*lac-proAB*), Δ(*mcrC-mrr*)102::Tn*10*(Tet^R)	Invitrogen
MG1655	F^-, *lam*$^-$, *rph*-1, *lacZ*$^+$	(Guyer et al, 1981)
MG1655Δ*dapA*	As MG1655, except *dapA*<>*cat*	This work
MGM100	As MG1655 except P_{BAD}*groESL*::Tn*5*(Kan^R)	(McLennan et al, 1998)
MGM100 *lacI*::Tn*10*	As MGM100 except *lacI*::Tn*10*(Tet^R)	This work
MGM100DapA·His	As MGM100, except containing a 6 x His tagged version of *dapA* linked to a Cmp^R cassette	This work
MGM100DapA·HSV	As MGM100, except containing a HSV tagged version of *dapA* linked to a Cmp^R cassette	This work
MGM100Δ*dapA*	As MGM100 except *dapA*<>*cat*	This work
MGM100Δ*dapA*λP_{dapA}-*lacZ*	As MGM100λP_{dapA}-*lacZ*, except *dapA*<>cat	This work

Strain	Genotype	Source/Reference
MGM100*lysA*λP_{dapA}-*lacZ*	As MGM100λP_{dapA}-*lacZ*, except *lysA*::*Tn*10(Tet^R)	This work
MGM100*thrA*λP_{dapA}-*lacZ*	As MGM100λP_{dapA}-*lacZ*, except *thrA*::*Tn*10(Tet^R)	This work
MGM100λDE3	As MGM100 except λDE3	This work
MGM100λP_{dapA}-*lacZ*	As MGM100, except λP_{dapA}-*lacZ*	N. McLennan
MGM100λRSPD72	As MGM100λP_{dapA}-*lacZ* except only 70 bases of P_{dapA} upstream of ATG used	This work
NK6056	Hfr, DE (gpt-lac)5, *lam*$^-$, *purC*80::Tn*10*(Tet^R), *relA*1, *spoT*1, *thi*-1	N. Kleckner
NL131	F$^-$, *thr*20, *leu*32, *proA*35, *codA*2, *lacY*1, *tsx*70, *gal*6, λ^S, λ$^-$, *rpsL*125, *xyl*7, *pyr*63, *groES*131^{ts}	N. McLennan
NM621	*Leu*, *proB*1, *hsdR*, *mcrA*$^-$, *mcrB*$^+$, *recD*, *supE*, *tsx*, *tonA*	(Whittaker et al, 1988)
TOP10	F, *mcrA*, Δ(*mrrhsdRMSmcrBC*), ϕ80*lacZ*ΔM15, Δ*lacX*74, *deoR*, *recA*1, *araD*139, Δ(*ara-leu*)7697, *galU*, *galK*, *rpsL* (str^R), *endA*1, *nupG*	(Grant et al, 1990)
TP8503	F$^-$, Δ(*lac-proB*), *leu*, *thi*, *supE*42, *fhuA*	Laboratory stock

Table 2.1.2 Yeast strains used in this study

Strain	Genotype	Source/Reference
S. cerevisiae EGY48	MATα, trp1, his3, ura3, leu2::6, LexAop-LEU2	MoBiTec

Table 2.1.3 Bacteriophages used in this study

Bacteriophage	Description	Source/Reference
P1	P1 K_C	Laboratory Stock
λDE3	Imm^{21}, Δnin5, Sam7, T7 polymerase gene controlled by the *lacUV5* promoter	(Studier et al, 1986)
λP_{sfiA}-*lacZ*	λ phage containing the *sulA* promoter (P_{sfiA}) transcriptionally fused to *lacZ*	(Huisman et al, 1983)
λP_{dapA}-*lacZ*	λ phage containing the *dapA* promoter (P_{dapA}) transcriptionally fused to *lacZ*	Laboratory stock
λRS45	Transcriptional promoter fusion vector with *lacZ* as reporter gene, imm^{21}, *ninR5*	(Simons et al, 1987)

Table 2.1.4 Plasmids used in this study

Plasmid	Description	Source/Reference
pBAD-18	Amp^R, 4.6 kb, *araC*, P_{BAD}, pBR322 replicon	(Guzman et al, 1995)
pBADLacZ	Amp^R, pBAD18 + 4.0kb *lacZ* insert	Laboratory stock
pBR322	Amp^R, Tet^R, pMB1 replicon	(Bolivar, 1978)
pBR325	Amp^R, Tet^R, Cmp^R, pMB1 replicon, ($oriT^+$)	(Bolivar, 1978)
pBRA6	Amp^R, pBR322 with 3.4 kb fragment containing *dapA*	Laboratory stock
pGEM-T	Amp^R, 3.0kb general purpose cloning vector	Promega
pGNG2	Amp^R, 6.7 kb, pBR ori, 2μm ori, URA3, promoter GFPuv expression under control of GAL1,10	MoBiTec
pGroEL	Cmp^R, pACYC + 2.1kb $groES^-$, $groEL^+$ insert	(Goloubinoff et al, 1989)

Plasmid	Description	Source/Reference
pJF118(EH)	Amp^R, ColE1 replicon, *tac* promoter and $lacI^Q$	(Furste et al, 1986)
pJF118DapA·His	As pJF118(EH), + 1.2 kb *dapA.His* cloned into the *Eco*RI and *Hin*dIII sites	This work
pJF118LacZ	As pJF118(EH) + 4 kb LacZ fragment cloned into *EcoR*I and *Hind*III sites	This work
pJG4-5	Amp^R, 6.4 kb, pUC ori, 2μm ori, TRP1, B42 under control of GAL1	MoBiTec
pPRO	Spc^R plasmid carrying *tetR* and $P_{lacI}{}^Q/lacI$	Clontech
pPROTet.E132DapA·His	Cmp^R, 3.2 kb plasmid, *dapA.His* expression under the control of $P_{Ltet0-1}$	This work
pRS551	Amp^R, Kan^R, 12.5 kb plasmid containing *lacZ* as a reporter	(Simons et al, 1987)
pRS551-PD72	As pRS551 + 70 bp of *dapA* promoter DNA ligated into the *Eco*RI and *Bam*HI sites	This work
pT7-5	Amp^R, 2.4 kb, Col E1 replicon, ϕ10 promoter	(Tabor et al, 1985)
pT7-5DapA·His	As pT7-5 + 1.2 kb *dapA.His* fragment.	This work
pUC18	General purpose cloning vector. 2.7 kb. Amp^R	(Yanisch-Perron et al, 1985)
pUCAT18	pUC18 + 1.45 kb *cat* gene from pBR322	Laboratory stock
pUCDapA	pUC18 + 3.4 kb fragment containing *dapA*	Laboratory Stock
pUCDapHis	Amp^R, Cmp^R, as pUCDapA except *dapA* C terminal tagged with 6 x HIS and *cat* gene present	This work
pUCDapHSV	Amp^R, Cmp^R, as pUCDapHis except *dapA* C terminal tagged with HSV	This work

Table 2.1.5 Growth media

L-broth	Difco bacto tryptone	10 g
	Difco bacto yeast extract	5 g
	NaCl	5 g
	pH to 7.2 with NaOH	
	Distilled water to 1 litre	
LB – NaCl	As L-broth, omitting NaCl	
L-agar	L-broth + 15 g Difco agar per litre	
LB top agar	L-broth + 6.5 g Difco agar per litre	
SOC broth	Bactotryptone	4 g
	Difco bacto yeast extract	1 g
	5 M NaCl	0.4 ml
	1 M $MgCl_2$	2 ml
	1 M KCl	0.5 ml
	1 M $MgSO_4$	2 ml
	glucose	0.72 g
	Distilled water to 200 ml	
Spizizen's minimal media	$(NH_4)_2SO_4$	10 g
	K_2HPO_4	70 g
	KH_4PO_4	30 g
	Sodium citrate·$2H_2O$	5 g
	$MgSO_4 \cdot 7H_2O$	1 g
	Distilled water to 1 litre	
VB minimal media	20x VB salts	50 ml
	20% carbon source	10 ml
	Thiamine HCl (1 mg/ml)	2 ml
	Supplements as required	
	Distilled water to 1 litre	
VB minimal agar	VB minimal media + 15 g difco agar	

20x VB salts	Citric acid	40 g
	KH_2PO_4	400 g
	$NaNH_4 \cdot HPO_4 \cdot H_2O$	70 g
	Distilled water to 1 litre	
	Store over 1 ml of chloroform	
YPD broth	Peptone	20 g
	Difco bacto yeast extract	10 g
	Glucose	20 g
	Distilled water to 1 litre	
YPD agar	YPD broth + 17 g Difco agar per litre	
DOB glucose	Yeast nitrogen base	1.7 g
	Ammonium sulphate	5 g
	Glucose	20 g
	Distilled water to 1 litre	
DOB galactose	As DOB glucose, omitting glucose and adding:	
	Galactose	20 g
	Raffinose	10 g
DOB agar	DOB glucose or galactose + 17 g Difco Agar per litre	

Table 2.1.6 Commonly used buffers

Phage buffer	Na_2HPO_4	7 g
	KH_2PO_4	3 g
	NaCl	5 g
	1 M $MgSO_4$	1 ml
	0.1 M $CaCl_2$	10 ml
	1% (w:v) gelatin solution	1 ml
	Distilled water to 1 litre	
Bacterial buffer	$MgSO_4 \cdot 7H_2O$	2 g
	Na_2HPO_4	7 g
	KH_2PO_4	3 g
	NaCl	4 g
	Distilled water to 1 litre	
TE buffer	10 mM Tris-HCl (pH 8.0)	
	1 mM EDTA (pH 8.0)	
TAE buffer	Working solution:	
	40 mM Tris-acetate	
	2 mM EDTA	
	Tris base	242 g
	Glacial acetic acid	57.1 ml
	0.5 M EDTA (pH 8.0)	100 ml
	Distilled water to 1 litre	
TBE buffer	Working solution:	
	89 mM Tris-borate	
	89 mM boric acid	
	Tris base	54 g
	Boric acid	27.5 g
	0.05 M EDTA (pH 8.0)	20 ml
	Distilled water to 1 litre	

2.1.3 Minimal medium supplements

Amino acid supplements were stored in stock solutions of pure amino acids at a concentration of between 2-10 mg/ml depending on the solubility of the particular amino acid. Sparingly soluble amino acids, such as tyrosine, were dissolved in 0.01M NaOH. The final concentration of the amino acids in the media was usually in the order of 20-100 μg/ml. If a rich minimal medium was required, vitamin free casamino acids (CAA) were used. The stock concentration of CAA was 100 mg/ml and the final concentration in the medium was typically 1-5 mg/ml. It should be noted that casamino acids lack tryptophan and this should therefore be added to the CAA medium if the bacterial strain to be used is auxotrophic for this amino acid.

Purines and pyrimidines were added to minimal media when required. Thymine and uracil were stored at a concentration of 2 mg/ml in water, and their final concentration in minimal medium was usually 20-40 μg/ml.

The only vitamin supplement used in this work was thiamine hydrochloride (vitamin B1). This was stored as a 1 mg/ml solution in water and its final concentration in minimal medium was 2 μg/ml.

2.1.4 Selection of antibiotic resistance

The routine concentrations for the antibiotics used in this work are shown in Table 2.1.6. Those antibiotics dissolved in water were filter sterilised using a 0.45 μm filter (Gelman).

Table 2.1.6 Antibiotic solutions

Antibiotic	Abbreviation	Solvent	Conc. of stock solution (mg/ml)	Final conc. in media (μg/ml)
Ampicillin	Amp	H_2O	100	50-100
Chloramphenicol	Cmp	Ethanol	20	25
Kanamycin sulphate	Kan	H_2O	25	25-50
Naladixic acid	Nal	0.1 M NaOH	20	20
Spectinomycin dihydrochloride	Spc	H_2O	50	25-50
Streptomycin sulphate	Str	H_2O	100	20
Tetracycline hydrochloride	Tet	50% ethanol	10	10

2.2 DNA techniques

2.2.1 Plasmid DNA preparation

Plasmid DNA was prepared using the Qiaprep spin miniprep kit (Qiagen) according to the manufacturers instructions.

2.2.2 Chromosomal DNA preparation

Chromosomal DNA was prepared using the AquaPure Genomic DNA Isolation kit (Biorad Laboratories) according to the manufacturers instructions.

2.2.3 DNA precipitation

DNA was precipitated from aqueous solution by: (i) adding 1/10 volume of 3M sodium acetate (pH 5.0) and 3 volumes of absolute ethanol, mixing thoroughly and leaving on ice for a minimum of ten minutes. This was then centrifuged in a microcentrifuge at 14000 x g for 15 minutes. The supernatant was discarded, and the pellet washed in 70% ethanol by vortexing. This was recentrifuged as above for 10 minutes, the supernatant again discarded and the pellet was dried under vacuum. The dried DNA pellet could then be resuspended in a suitable volume of TE buffer (with added RNase (20μg/ml) if required). (ii) Instead of 3 volumes of absolute ethanol, 1 volume of isopropanol could be used. This had the advantage of keeping the total volume smaller, and was therefore the preferred method. After isopropanol precipitation and centrifugation, the pellet was washed with 70% ethanol as above.

2.2.4 Determination of DNA concentrations

DNA concentrations were determined by measuring the absorption of diluted DNA solutions at 260 nm. For double stranded DNA, an OD_{260} value of 1.0 represents a DNA concentration of 50 μg/ml, and for single stranded DNA the same value represents a DNA concentration of 40 μg/ml. DNA purity can be determined by measuring absorption at 260 and 280 nm. Protein-free double stranded DNA should give a 260/280 ratio close to 1.8, and single stranded DNA should give a ratio nearer 2.0.

2.2.5 Digestion of DNA with restriction endonucleases

Endonuclease cutting of DNA was typically performed in volumes of between 10 and 100 μl. These contained the requisite amount of DNA (usually 1-10 μg) and the appropriate Boehringer Mannheim restriction buffer at a 1x concentration. The restriction enzyme was usually present in a two - fivefold excess, i.e. 2-5 units per microgram of DNA. The digests were made up to their final volume using distilled water. The complete restriction digests were incubated at the recommended temperature (usually 37°C) for 1-3 hours. The products of the reaction were either directly analysed by agarose gel electrophoresis, or phenol extracted, ethanol precipitated and dissolved in a suitable volume of TE buffer for further manipulation.

2.2.6 Partial digestion of DNA

For partial digestion of DNA, ten two-fold serial dilutions of restriction enzyme were added to fixed amounts of DNA, with 0.5 units of enzyme per microgram of DNA representing the highest enzyme:DNA ratio. The digests were incubated at the appropriate temperature for 1 hour and terminated by addition of tracking dye (Section 2.2.9). The products of the reactions could then be analysed by agarose gel electrophoresis.

2.2.7 Ligation of DNA

Ligations of DNA were typically performed in a final volume of 10 µl. These contained 1x Boehringer Mannheim ligation buffer, T4 DNA ligase and between 0.5 and 1 µg total DNA with insert DNA in a 2 to 20 molar excess over the vector DNA. Ligase (0.2 units) was used for the ligation of cohesive DNA termini, and 1 unit of the enzyme was used for the ligation of blunt ended molecules. These reactions were incubated for at least 12 hours at 16°C. Between 5 and 10 µl of the reaction mixture was then used to transform competent cells of an appropriate strain of *E. coli*.

2.2.8 'Filling in' of recessed 3' termini

DNA polymerase I Klenow fragment was used to fill-in the recessed 3' termini generated by various restriction enzymes to give blunt ended DNA molecules. Reactions were performed in a final volume of 20 µl containing 1 µg DNA, 1x Klenow buffer, all four dNTPs, each at a concentration of 20 µM and 2 units of Klenow enzyme. The reactions were incubated at 16°C for 45 minutes. The reaction was stopped and the unincorporated nucleotides removed by increasing the reaction volume to 200 µl with TE buffer, phenol extracting and ethanol precipitating the DNA.

2.2.9 Agarose gel electrophoresis

Agarose gel electrophoretic analysis of DNA was always performed using TAE buffer. The gels were made by melting the appropriate amount of agarose in 1x TAE buffer using a microwave oven. For fragments of 300 bp to 1.5 Kb 1.5% (w:v) agarose was used, between 1.5 Kb and 4 Kb 1% agarose and above 4 Kb 0.8% (w:v) agarose. The molten agarose was poured into the gel tray, the comb inserted and the gels left to solidify for 30 minutes. Once set, the comb was removed from the gel and he DNA samples containing 1x tracking dye (6x tracking dye is 0.25% bromophenol blue, 0.25% xylene cyanol and 40 % (w:v) sucrose in distilled water) were loaded into the wells. Gels were run in Pharmacia gel electrophoresis tanks with their surfaces only just immersed in 1x TAE buffer. Electrophoresis was either performed overnight with a constant current of 25 mA or in 2-3 hours with a constant current of 80 mA. After completion of electrophoresis, gels were stained in water containing 2 μg/ml ethidium bromide for about 1 hour with constant shaking, and subsequently destained in fresh water for 30 minutes. The gels could then be photographed using Polaroid film and UV transillumination.

2.2.10 Isolation of DNA from agarose gel slices

To isolate DNA from agarose gel slices a low melting temperature agarose was used to separate the DNA fragment of interest and the DNA was then extracted from the gel using β-Agarase I (NEB). β-Agarase cleaves the agarose subunit, unsubstituted neoagarobiose (3,6-anhydro-α-L-galactopyranosyl-1-3-D-galactose) to neoagaro-oligosaccharides (Yaphe, 1957).

The DNA was run on a low melting temperature agarose gel. The agarose used depended on the size of the DNA fragment of interest, NuSieve GTG agarose was used for fragments less than 1 Kb and SeaPlaque GTG agarose was used for DNA fragments larger than 1 Kb (both NuSieve and SeaPlaque GTG agarose are available from FMC Bioproducts). The agarose gel was run using the protocol outlined in section 2.2.9, the appropriate DNA band was located using UV transillumination and cut out using a clean scalpel in as small a volume of agarose as possible. Gel slices were transferred to Eppendorf tubes, the weight determined and 1/10 volume of 10x β-Agarase buffer added. The gel slice and buffer were then placed at 65°C for 10 minutes to melt the agarose and then cooled to 40°C. 2 units of β-Agarase I were then added and the tube was incubated at 40°C for 1 hour. After the

1-hour incubation the DNA could be precipitated, leaving the carbohydrates in solution, using the protocol outlined in section 2.2.3.

10x β-Agarase buffer:	100 mM Bis-Tris-HCl
	10 mM EDTA
	pH 6.5

2.2.11 Labelling DNA fragments using T4 Polynucleotide Kinase

The DNA to be labelled was either a PCR product, which had been ethanol precipitated, or a commercially produced primer. 50 pmol of 5' termini were used in a 50 µl reaction, this would correspond to 50 pmol of single stranded DNA or 25 pmole double stranded DNA. A typical reaction would be as follows:

DNA	25 or 50 pmol
10x T4 PNK buffer	5 µl
[γ-^{32}P] ATP	5 µl (5 µCi)
T4 Polynucleotide Kinase	20 units
Water to 50 µl	

The reaction was incubated for 30 minutes at 37°C and the unincorporated [γ-^{32}P] ATP removed by passing the reaction mixture through a MicroSpin G-25 column (Amersham Pharmacia Biotech).

2.2.12 DNA sequencing technique

DNA sequencing was performed using the Pharmacia Sequenase quick-denature plasmid sequencing kit. The kit is based on the chain-terminating dideoxynucleotide sequencing method developed by Sanger *et al* (Sanger et al, 1977). In the original procedure, primer extension was catalysed by the Klenow fragment of *E. coli* DNA polymerase I. In the kit, however, the Klenow enzyme has been replaced by Sequenase DNA polymerase (Tabor et al, 1987), which has the advantage of creating longer chain-terminated fragments with a more even distribution of label between fragments. The major practical difference in using Sequenase DNA polymerase is that the primer extension reactions are performed in two stages, a labelling reaction and a termination reaction. Two stages are required

because the enzyme uses dideoxynucleotides very readily, and therefore in order to allow the synthesis of long chain-terminated fragments, dideoxynucleotides are excluded from the first stage of the reaction whilst being added for the second.

(i) Annealing of primer to double stranded template. The DNA templates used in the sequencing reaction were usually double-stranded plasmid or chromosomal DNA, purified as mentioned earlier. The concentration of the template was adjusted so that 5 µl contained 0.5-3 µg DNA, and 1 µl of primer contained 2-4 pmol. To denature the double stranded template the following was added to an Eppendorf tube:

Template DNA	5 µl
1 M NaOH	2 µl
Water	3 µl
Primer	1 µl

The tube was then vortexed briefly and allowed to incubate at 37°C for 10 minutes. Plasmid DNA will denature (at any temperature) when exposed to pH 13. The method outlined in this protocol requires that the alkali used to denature the DNA be precisely neutralized by the addition of HCl.

The tube was then put on ice and the following added:

1 M HCl	2 µl
Plasmid reaction buffer	2 µl

The tube was then capped and returned to the 37°C water bath for 10 minutes to promote annealing of the primer to the single stranded DNA.

(ii) Sequencing reaction. For each template to be sequenced, four Eppendorf tubes were labelled 'A', 'C', 'G' and 'T' respectively and 2.5 µl of the corresponding dideoxynucleotide mix added to each tube. To the tube containing the annealed template and primer the following was added:

Annealed template/primer	15 µl
0.1 M DTT	1 µl
Labelling mix	2 µl
[α-^{32}P] dATP	1 µl (10 µCi)
Sequenase plasmid sequencing formulation	2 µl

The labelling reaction was incubated at room temperature for 5 minutes. While this was proceeding the previously dispensed sequencing mixes were incubated at 37°C for a minimum of 1 minute. After the 5 minute incubation of the labelling reaction, 4.5 µl was added to each of the prewarmed sequencing mixes and

returned to the 37°C water bath for a further 5 minutes to allow the chain-termination reaction to occur. Finally, 5 µl of Stop Solution was added to each reaction, which could then be stored at –20°C until required for electrophoresis. When the samples were needed for loading onto the sequencing gel they were heated to 80°C for 2 minutes to denature the DNA. Immediately after this incubation 1.5-2.5 µl of each sample was loaded onto the gel.

(iii) DNA sequencing gel electrophoresis. DNA sequencing was performed using a 30 x 40 cm BRL sequencing apparatus. The glass sequencing gel plates were thoroughly cleaned with ethanol and chloroform assembled using 0.2 mm spacers and taped together carefully to minimise the possibility of leakage.

The gel was prepared by adding together the following:

Bis-acrylamide (40% w:v)	30 ml
Urea	43 g
10x TBE	10 ml
Water	20 ml

This was allowed to dissolve with the aid of magnetic stirring. Once dissolved 1 ml of 10% ammonium persulphate solution was added followed by 35 µl of TEMED. This was stirred slowly for a few seconds and was then poured between the sequencing plates. The flat edge of a 60-well shark-tooth comb was pushed between the plates to layer the top of the gel. Clingfilm was wrapped around the exposed areas of the gel and each edge of the gel was clamped with bulldog clips. The gel was then set aside for at least 1 hour to allow polymerisation. Once set, the bulldog clips, tape and comb were removed and distilled water was squirted along the top of the gel. The shark-tooth comb was then replaced with the points downwards just touching the surface of the gel. The gel was then clamped into the sequencing apparatus and 1x TBE solution poured into the top and bottom reservoirs. The gel was pre-run at about 50 W (~1500 V) for 1 hour. After this the gel was ready to be loaded with the sequencing reactions. The samples were loaded immediately after denaturing the DNA (see above). The gel was then electrophoresed at 50 W until the blue dye front ran off the end of the gel. Once electrophoresis was complete the glass plates were removed from the apparatus and the top plate very carefully removed. The bottom plate (with gel attached) was placed in a fixing bath containing 40% methanol and 10% acetic acid in water for 20 minutes. The plate and gel were then removed and two damp sheets of blotting paper placed on top of the gel followed by two dry sheets of blotting paper. Even pressure was applied and the papers were peeled off the glass plates taking the gel with them. The gel and paper sandwich was

then dried in a vacuum gel drier for 1 hour at 80°C. When dry, the gel was placed in an autoradiography cassette and allowed to develop at ambient temperature. In most cases a good signal was achieved after 24 hours.

2.2.13 Amplification of DNA using the polymerase chain reaction

The polymerase chain reaction (PCR) was performed to confirm gene replacements and to create engineered molecules for cloning. Primers were usually between 17-24 bases in length and were obtained commercially (Genosys). For all PCR reactions a commercial enzyme mix containing thermosTable *Taq* DNA polymerase and a proofreading polymerase was used (Roche). A typical reaction mixture might be:

Enzyme mix	1 µl
10x buffer	5 µl
Primer 1	300 nM
Primer 2	300 nM
dNTP's	350 µM each
Plasmid template	10 ng
Chromosomal template	500 ng
Distilled water to 50 µl	

The melting temperature used was 94°C for 2 minutes. Annealing temperature varied depending on which primers were being used. The annealing temperature (T_m) was gauged by the following formula.

$$69.3 + 0.41(G + C)\% - 650/L = T_m$$

where L is the length of the oligonucleotide in nucleotides and (G + C)% represents the percentage guanine and cytosine content (Maniatis et al, 1983). The extension temperature used was always 68°C. As a rule of thumb, 30 seconds per Kb of extension was allowed.

2.2.14 Site directed mutagenesis of plasmid DNA

Site directed mutagenesis of plasmid DNA was performed using the QuikChange Site Directed Mutagenesis Kit (Stratagene). Mutagenesis was performed as per the manufacturers instructions. A brief summary of the protocol is described here; the QuikChange site-directed mutagenesis kit is used to make point mutations, switch amino acids, and to delete or insert single or multiple amino acids. The basic procedure utilises a supercoiled double-stranded DNA (dsDNA) vector with an insert of interest and two synthetic oligonucleotide primers containing the desired mutation. The oligonucleotide primers, each complementary to opposite strands of the vector, are extended during temperature cycling by using *PfuTurbo* DNA polymerase. Incorporation of the oligonucleotide primers generates a mutated plasmid containing staggered nicks. Following temperature cycling, the product is treated with *Dpn*I. The *Dpn*I endonuclease is specific for methylated and hemimethylated DNA and is used to digest the parental DNA template and to select for mutation-containing synthesised DNA (Nelson et al, 1992). DNA isolated from almost all *E. coli* strains is dam methylated and therefore susceptible to *Dpn*I digestion. The nicked vector DNA incorporating the desired mutations is then transformed into XL-1 Blue supercompetent cells, where the nicks in the mutated plasmid are repaired. The small amount of starting DNA template required to perform this method, the high fidelity of the *PfuTurbo* DNA polymerase, and the low number of thermal cycles all contribute to the high mutation efficiency and decreased potential for random mutations during the reaction.

2.3 Bacterial Techniques

2.3.1 Preparation and transformation of competent cells

For transformation of plasmids and ligation mixtures the following procedure was followed. 5 ml of medium containing the appropriate selective antibiotics was inoculated with a single colony and incubated with shaking at the appropriate temperature overnight. 200 µl of the overnight culture was used to inoculate 20 ml of L-Broth, with appropriate selection, and incubated at a suitable temperature until the OD_{600} reached 0.6. The culture was then transferred to a sterile Universal bottle and placed on ice for 10 minutes. The chilled culture was centrifuged in a bench centrifuge at 1500 x g for 5 minutes. The supernatant was discarded and the universal

was inverted to drain off excess media. The cells were resuspended in 10 ml ice-cold 0.1 M $MgCl_2$/0.1 M $RbCl_2$ (9:1 v:v). The cell suspension was allowed to sit on ice for 10 minutes and was then re-centrifuged at 1500 x g for 5 minutes. The supernatant was removed and the cells then resuspended in 1 ml 0.1 M $CaCl_2$/0.1 M $RbCl_2$ (9:1 v:v). The cells were stored on ice in a 4°C cold room overnight and were ready for use the next day. Immediately before use 1.5 µl dimethylsulphoxide (DMSO) was added per 100 µl of competent cells. 5 µl of plasmid DNA or 10 µl of ligation reaction were added to 100 µl of competent cell, mixed gently and placed on ice for 30 minutes. The mixture was then transferred to a 42°C waterbath for 45 seconds and then placed on ice for 5 minutes to recover from the heat shock. 250 µl of L-broth (containing 0.2 mM of an appropriate carbon source) was then added to the cells and the culture incubated on a rotating wheel at a suitable temperature for 1 hour, to allow the expression of any antibiotic selection marker and then 350 µl of the culture was plated on selective media.

2.3.2 Frozen storage of bacterial strains

It was found that strains of *E. coli* could be conveniently stored at -70°C without suffering a dramatic loss of viability; this included strains harbouring plasmids that might otherwise be lost. A fresh 5 ml overnight culture was prepared with antibiotic selection if required. This was centrifuged at 1500 x g for 10-15 minutes, the supernatant discarded and the cells resuspended in 1/10 the original volume of frozen storage buffer. The cells were then left on ice for a couple of hours before storing at -70°C.

Frozen storage buffer: 50%Bacterial Buffer
50% Glycerol (v:v)

2.4 Bacteriophage techniques

2.4.1 Preparation of λ plate lysates

Cells were grown in L-broth + 20 mM $MgSO_4$ and maltose at 0.2% until mid-log phase. 200 µl aliquots were then mixed with 10^6 phage, incubated at 37°C for 5 minutes, and 3 ml of L-top agar containing 20 mM $MgSO_4$ and 0.2% maltose added.

This was poured onto an L-agar plate, left to set and incubated at 37°C overnight or until visible lysis occurred. 5 ml of phage buffer was then added to the plate and the layer of top agar scraped off into a sterile 250 ml beaker. A few drops of chloroform were added and the beaker was incubated at 30°C with gentle swirling for 20 minutes. The contents of the beaker was poured into a universal bottle and centrifuged at 1500 x g for 10 minutes. The supernatant was transferred to a fresh half ounce bottle and stored over a few drops of chloroform at 4°C.

2.4.2 Preparation and selection of λ lysogens

A lawn of the bacterial strain to be lysogenised was made by mixing 0.2 ml of a mid-log phase L-broth culture and 3 ml of L-top agar with 20 mM $MgSO_4$ and 0.2% (w:v) maltose added. To this, approximately 200 λ phage particles were added, and the mixture poured onto a fresh L-agar plate. Once set the plate was incubated overnight at 37°C. This resulted in the appearance of isolated λ plaques. The centre of a plaque was then touched with a sterile toothpick and streaked out onto another plate, which again was incubated overnight. The resulting single colonies could now be tested for the presence of the λ phage. A lysogenised bacterium would be resistant to lysis by phages with the same immunity as the one used to lysogenise the strain, but will be sensitive to λ phages that are virulent, or carrying a different immunity. This was initially done as a cross-streak on an L-agar plate, and positives from this test were checked more fully by trying to make isolated plaques in a similar way to that mentioned above. Strains that seemed to be lysogens were then finally tested by examining supernatants from L-broth cultures to see if they were producing the correct phage upon induction.

2.4.3 UV induction of λ lysogens

Lysogenic bacteria were grown in L-broth + 20 mM $MgSO_4$ at 37°C with vigorous agitation until an OD_{600} of 0.4 was achieved. The cells were harvested by centrifugation and resuspended in 7 ml of 20 mM $MgSO_4$. This was transferred to a sterile glass petri dish and the cells were exposed to 600 ergs/mm^2/sec of UV irradiation before being diluted fivefold in fresh L-broth + 20 mM $MgSO_4$. The culture was grown at 37°C with vigorous shaking until lysis occurred. A few drops of chloroform were added, and the lysate clarified by centrifugation prior to titration.

2.4.4 Preparation of phage P1 lysates

Preparation of phage P1 lysates were as for phage λ except that 10^6 phage were added to 1 ml of late-log phase cells and this was incubated at 37°C for 30 minutes prior to the addition of the top agar. The maltose was omitted and the $MgSO_4$ was replaced with 2.5 mM $CaCL_2$. The phage buffer was also replaced with the same volume of L-broth + 2.5 mM $CaCl_2$.

2.4.5 Phage P1-mediated transduction

The recipient strain of *E. coli* was grown up to late-log phase in L-broth. The cells were harvested by centrifugation and the bacterial pellet resuspended in 1/10 the volume of L-broth + 2.5 mM $CaCl_2$. 100 µl aliquots of this 10x culture were mixed with either 0.1 ml of phage P1 stock or 0.1 ml of a 10x dilution of the phage stock. These were incubated at 37°C for 15 minutes. If prototrophic transductants were to be selected, 0.4 ml of phage buffer was added to 0.2 ml aliquots of the cells and plated on the appropriate minimal media agar plates. If the selection was the acquisition of antibiotic resistance then 1 ml of phage buffer was added to the cells, the cells were centrifuged and resuspended in 0.6 ml of L-broth + 20% glucose. This was incubated at 37°C for 1 hour to allow expression of the antibiotic resistance and then 0.2 ml aliquots were plated out on L-broth agar plates containing the appropriate antibiotics. Plates were incubated at a suitable temperature until colonies appeared.

2.5 Protein Techniques

2.5.1 Polyacrylamide gel electrophoresis of proteins

E. coli proteins were routinely separated using SDS-polyacrylamide gel electrophoresis with a discontinuous buffer system (Laemmli, 1970). In most cases a 10% resolving gel and a 4% stacking gel were employed. For the rapid analysis of proteins a 'Mighty-Small' mini-gel apparatus (SE 250) manufactured by Hoefer Scientific instruments was used. This used 10 x 5 cm gels, which could be electrophoresed in about 45 minutes. For better quality gels a larger apparatus (SE

600) made by the same company was used; in this case the gels measured 16 x 18 cm. The composition of typical gel solutions used were as follows:

10% resolving gel	acrylamide solution (40%)	2.5 ml
	4 x resolving gel buffer	2.5 ml
	10% SDS	0.1 ml
	10% ammonium persulphate	0.1 ml
	distilled water	4.8 ml
	TEMED	10 µl
4% stacking gel	acrylamide solution (40%)	0.5 ml
	4 x stacking gel buffer	1.25 ml
	10% SDS	0.05 ml
	10% ammonium persulphate	0.05 ml
	distilled water	3.15 ml
	TEMED	5 µl

These solutions were made up immediately prior to use, with the ammonium persulphate solution and TEMED being added last. The resolving gel solution was pipetted between the glass plates separated by 0.75 mm spacers, enough room being left for the stacking gel. For the mini-gel apparatus the depth of the stacking gel between the bottom of the comb and the resolving gel was about 1 cm, and for the larger gel, was about 2.5 cm. Once the resolving gel had been poured it was layered with isobutanol saturated with 1x resolving-gel buffer and allowed to polymerise for 20 minutes. The isobutanol was then discarded and the top of the gel was washed with distilled water. The stacking-gel solution was poured on top of the resolving gel, the comb inserted and polymerisation allowed to occur. The comb was then removed and the wells washed out with 1x reservoir buffer, which was also used to fill up the buffer chambers of the apparatus. The sample could be loaded onto the gel at this stage.

Samples were mixed 1:1 with 2x PAGE-loading buffer, boiled for 3 minutes, and briefly spun in a microcentrifuge prior to loading. About 5-10 µl of sample per lane could be loaded onto the mini-gel and 10-30 µl on the larger apparatus. Gels were typically electrophoresed at a constant current of 15 mA per gel for the mini-gels and 30 mA per gel for the larger format gels, until the bromophenol blue dye front had reached the bottom of the gel. Once electrophoresis was complete the glass plates were removed from the apparatus, separated carefully using a plastic wedge and the gel placed in staining solution for 1 hour with gentle agitation. Gels were

transferred into destaining solution and left for between 2-24 hours, or until the bands of interest could be seen clearly. For preservation, the stained/destained gel was soaked in destaining solution plus 5% glycerol for 30 minutes and dried onto blotting paper using a vacuum gel drier at 80°C for about 1 hour.

Solutions used in SDS-PAGE:

Stock acrylamide	acrylamide	37 g
	NN' methyl bis-acrylamide	1 g
	Made up to 100 ml with distilled water	
	Filtered and stored at 4°C	
4x stacking-gel buffer	Tris base	15.25 g
	made up to 250 ml with distilled water	
	pH adjusted to 6.8, autoclaved	
4x resolving-gel buffer	Tris base	45.5 g
	made up to 250 ml with distilled water	
	pH adjusted to 8.8, autoclaved	
10x reservoir buffer	Tris base	30.2 g
	Glycine	144 g
	Made up to 1 l with distilled water	
	0.1% SDS was added to the 1x buffer before use	
2x PAGE loading buffer	4x stacking-gel buffer	125 µl
	10% SDS	300 µl
	50% glycerol	200 µl
	2-mercaptoethanol	50 µl
	0.1% bromophenol blue	200 µl
	Distilled water	125 µl
Staining solution	9% (v:v) acetic acid	
	45% (v:v) methanol	
	0.1% (w:v) Coomassie brilliant blue	
Destaining solution	7% (v:v) acetic acid	
	5% (v:v) methanol	

Native (non-denaturing) polyacrylamide gels could be made by omitting the SDS from the stacking and resolving gel solutions and also from the running buffer. Also, during sample preparation, the cells are resuspended 1:1 in Native loading buffer and rather than being boiled for 3 minutes, are lysed by sonication.

5x Native loading buffer	1 M Tris-HCl (pH 6.8)	3.1 ml
	Glycerol	5 ml
	1% Bromophenol Blue	0.5 ml
	Distilled water	1.4 ml

2.5.2 Two dimensional polyacrylamide gel electrophoresis

Complex protein mixtures extracted from *E. coli* cells were separated by two-dimensional electrophoresis (2-D electrophoresis) (O'Farrell, 1975) using the IPGphor isoelectric focusing system (Amersham Pharmacia biotech). This technique sorts proteins according to two independent properties in two discrete steps: the first-dimension step, isoelectric focusing (IEF), separates proteins according to their isoelectric points (pI); the second dimension step, SDS-polyacrylamide gel electrophoresis (SDS-PAGE), separates proteins according to their molecular size. Each spot on the resulting two-dimensional array corresponds to a single protein species in the sample. Thousands of different proteins can thus be separated, and information such as the protein pI, the apparent molecular size, and the amount of each protein is obtained.

1 ml of culture (OD_{600} ~ 0.2) was placed into an Eppendorf tube and centrifuged to pellet the cells, the supernatant was removed and the cells resuspended in 50 µl of lysis solution. The cells were then lysed by freeze thaw lysis (Lenstra et al, 1983) and centrifuged again to pellet any insoluble material. 45 µl of the protein solution was then mixed with 205 µl of rehydration solution and the entire volume pipetted carefully into a 13 cm ceramic strip holder. The protective cover of the IPG strip (usually pH 3-10NL, 13 cm) was removed and the strip lowered, gel side down, into the ceramic holder. The strip was positioned so that the anodic end of the strip was directed towards the pointed end of the strip holder and gel contacted the strip holder electrodes at each end. IPG cover fluid (neutral mineral oil) was then applied to prevent evaporation and urea crystallisation. The cover was placed onto the ceramic holder and the holder was then placed onto the IPGphor. The IPGphor was

set at 20°C, 50 μA per IPG strip and the following isoelectric focusing program was always used:

Step	Voltage (V)	Duration (hr:min)	Gradient type
rehydration	20	12:00	Step-n-hold
1	500	1:00	Step-n-hold
2	1000	1:00	Step-n-hold
3	8000	2:00	Step-n-hold

A low voltage was applied during the rehydration step as it was found to give sharper focusing and also prevented protease activity during this period.

After isoelectric focusing the strips were removed from the ceramic holders and placed into individual tubes, 10 ml of equilibration solution 1 was added, the tube was capped, placed on a rocker and allowed to equilibrate with gentle agitation for 15 minutes at room temperature. The strip was then removed to a new tube, 10 ml of equilibration solution 2 was added, and the tube was capped, placed on a rocker and again allowed to equilibrate for 15 minutes at room temperature with gentle agitation.

A 10% SDS-PAGE gel was poured using the Hoefer SE600 apparatus as per the protocol for polyacrylamide gel electrophoresis, except the stacking gel was omitted and 1.5mm spacers were used. After the gel had polymerised, the focused IPG strip was dipped into 1 x reservoir buffer to lubricate it. The strip was then positioned between the plates on the surface of the second-dimension gel with the plastic backing against one of the glass plates. The strip was then pressed down with a thin ruler so that the entire lower edge of the IPG strip is in contact with the top surface of the slab gel, ensuring that no air bubbles were trapped between the IPG strip and the surface of the slab gel. Molecular weight markers were applied to a paper IEF sample application piece in a volume of 15 to 20 μl and applied to the top surface of the gel next to one end of the IPG strip.

1.5 ml of molten agarose sealing solution (no hotter than 50°C) was pipetted onto the top of the gel to hold the strip in place. The gel was then electrophoresed using the conditions given in polyacrylamide gel electrophoresis.

Lysis solution	Urea	19.2 g
	CHAPS	1.6 g
	Tris Base	0.194 g
	Water	to 40 ml

Rehydration solution	Urea	12 g
	CHAPS	0.5 g
	IPG buffer (same pH range as IPG strip)	125 µl
	Bromophenol blue	a few grains
	Water	to 25 ml
Equilibration solution 1	Tris-Cl pH 8.8 (1.5M)	6.7 ml
	Urea	72.07 g
	Glycerol (87% v:v)	69 ml
	SDS	4 g
	Water	to 200 ml
	Just prior to use add 100 mg DTT to 10ml of buffer	
Equilibration solution 2	As per equilibration solution 1, except replace DTT with 250 mg iodoacetamide	
Agarose sealing Solution	Agarose	0.5 g
	Bromophenol blue	a few grains
	1 x reservoir buffer	to 100 ml
	Heat all ingredients until agarose completely dissolves. Dispense 2 ml aliquots and allow to cool.	

2.5.3 Western blotting procedures

Western blotting (the transfer of proteins from an acrylamide gel to a membrane) was performed using a BioRad electrophoretic transfer cell. The previously run gel had the stacking gel removed and was placed on top of several sheets of blotting paper saturated in 1x reservoir buffer without SDS (see section 2.5.1). A sheet of Hybond PVDF membrane saturated in the same buffer was placed on top of the gel followed by several more sheets of saturated blotting paper. The whole 'sandwich' was secured in the transfer apparatus and transfer was allowed to proceed at 1 amp for 1 hour. After this the PVDF was removed from the apparatus and placed in blocking solution for 1 hour at room temperature or overnight at 4°C. The membrane was briefly rinsed twice with PBS-T (PBS + 0.1% (v:v) Tween-20)

and then washed with an excess volume of PBS-T for 5 minutes. The membrane was now ready for the application of the primary antibody. The primary antibody was diluted in 15 ml of PBS-T to give optimal results, and the membrane incubated in this solution, with constant agitation, at room temperature for 1 hour. The membrane was then washed for 2 x 10 minutes in PBS-T and incubated with 15 ml of the appropriate dilution of the secondary antibody (mouse/rabbit HRP-linked anti primary IGg from sheep/donkey, Amersham Pharmacia) at room temperature, with constant agitation, for 1 hour. The membrane was finally washed for 3 x 10 minutes with PBS-T and was now ready for the chemiluminescent detection of the protein of interest using the ECL Plus kit (Amersham Pharmacia).

PBS:	Di-Sodium hydrogen orthophosphate	11.5 g
	Sodium dihydrogen orthophosphate	2.96 g
	NaCl	5.84 g
	Distilled water to 1L	

2.5.4 Determination of protein concentration

Protein concentration was determined using the Bio-Rad Protein Assay kit. The Bio-Rad protein assay is a dye-binding assay based on the differential colour change of a dye in response to various concentrations of proteins (Bradford, 1976). A protein standard (bovine gamma globulin) was diluted to give samples of 0.2-1.4 mg/ml and 100 µl of these standards were added to 5 ml of 1x dye reagent and mixed by inversion several times. An amount of the protein solution of unknown concentration was also added to 5 ml of 1x dye reagent and mixed by inversion. The tubes were left for 10 minutes and then the OD_{595} was measured using 1x dye reagent as a blank. The OD_{595} was plotted versus concentration of the standards, and the unknown concentration determined from the standard curve.

2.5.5 Purification of His-tagged proteins

His-tagged proteins were purified using His·Bind Kits (Novagen). For small amounts of protein (≤ 2.0 mg) His·Bind Quick 900 cartridges were used and for larger amounts of protein His·Bind resin was used (His·Bind resin can purify 20 mg of protein per 2.5 ml of resin).

His·Bind resin allows rapid one step purification of proteins containing a His-tag sequence by metal chelation chromatography. The His-tag sequence binds to Ni^{2+} cations, which are immobilised on the His·Bind resin using the charge buffer supplied in the His·Bind buffer kit. After unbound proteins are washed away, the target protein is recovered by elution with imidazole or EDTA. The His·Bind resin can be regenerated and reused many times.

His·Bind Quick 900 cartridges are based on a large diameter cellulose matrix that has flow rates 5-50 times faster than agarose resins, and yet maintains similar binding capacity. His·Bind Quick 900 cartridges are packed with pre-charged resin (Ni^{2+}) and are designed for operation with a syringe.

2.5.6 β-Galactosidase assays

The method used to determine the β-galactosidase activity from a promoter-*lacZ* clone, was as described by Miller (Miller, 1972). 1 ml of culture was taken and the OD600 measured; from this 0.5 ml of culture was added to 0.5 ml of Z buffer. If the promoter activity was known to be high then 0.1 ml of culture was added to 0.9 ml of Z buffer. The cells were permeabilised by the addition of 50 µl of chloroform and the samples vortexed for 30 seconds. The samples were stored at 4°C until all sampling had been performed.

To assay, 200 µl of 4mg/ml ONPG (o-nitrophenyl-β-D-galactosidase) in 0.1M sodium phosphate buffer (ph 7.5) was added to each sample and vortexed. A control was prepared from 0.5 ml of the culture medium used and 0.5 ml of Z buffer. The samples were then incubated at 30°C until sufficient yellow colour (o-nitrophenol) had appeared such that when 0.5 ml of 1M Na_2CO_3 was added to stop the reaction, the absorbance at 420 nm was between 0.1 and 2.0. Standard curves constructed in this laboratory by T. Paterson had demonstrated that the absorbance measurements are linear within this range. Samples were then measured at both 420 nm (o-nitrophenol) and 550 nm (to correct for light scattering by cell debris), using the control tube as a blank.

The specific β-galactosidase activity of the sample was calculated by the following equation and was expressed as Miller Units (MU):

$$\frac{(OD_{420} - 1.75 \times OD_{550})}{OD_{600} \times V \times T} \times 1000$$

Where T represents the incubation time at 30°C in minutes and V represents the volume of the sample in millilitres.

Z Buffer (per litre):	$Na_2HPO_4 \cdot 7H_2O$	0.06 M
	$NaH_2PO_4 \cdot H_2O$	0.04 M
	β-mercaptoethanol	0.05 M
	KCl	0.01 M
	$MgSO_4 \cdot 7H_2O$	0.001 M
	SDS	0.005%

2.5.7 Synthesis of aspartic semialdehyde

L-aspartic-β-semialdehyde was prepared as described by Black and Wright (Black et al, 1955). Briefly; 2.3 g (20 mmoles) of L-allylglycine were dissolved in 20 ml of 1 M HCl. Ozone was passed through a fritted glass plug into this solution, held at 0°C for 100 minutes. Before use the solution (which was ~ 95% L-aspartic-β-semialdehyde) must be neutralised with an equal volume of 1 M $KHCO_3$.

2.5.8 DHDPS activity assays

To determine DapA activity in *E. coli* two methods were used. The first assay utilises an absorption change at 270 nm and is not suitable for complex mixtures and was therefore only used when assaying purified DapA protein. The second assay utilises a colour change when o-aminobenzaldehyde is present, the advantage of this assay is its specificity, which makes it possible to employ it for the analysis of crude extracts (both assays are described in Yugari et al, 1965). A typical assay solution might contain:

270 μm assay	DapA solution	1 μl
	Imidazole pH 7.4 (1M)	100 μl
	L-aspartic semialdehyde (1M)	5 μl
	$KHCO_3$ (1M)	5 μl
	Sodium pyruvate (50 mM)	100 μl
	Water	789 μl
	Measure absorbance change at 270 nm.	

o-aminobenzaldehyde assay	Crude cell extract	100 µl
	Tris-Cl pH 7.4 (1M)	100 µl
	L-aspartic semialdehyde (1M)	5 µl
	$KHCO_3$ (1M)	5 µl
	Sodium pryuvate (50mM)	200 µl
	o-aminobenzaldehyde solution ($0.25\ mg\ ml^{-1}$)	2 µl
	Water	588 µl
	Measure absorbance change at 540 nm.	

In both assays, the rate of absorbance change is proportional to the activity of the enzyme. With the o-aminobenzaldehyde assay there is a lag before the reaction shows a uniform increase in colour with time, however the rate of the linear portion of the reaction is directly proportional to enzyme activity (Fig. 2.5.1).

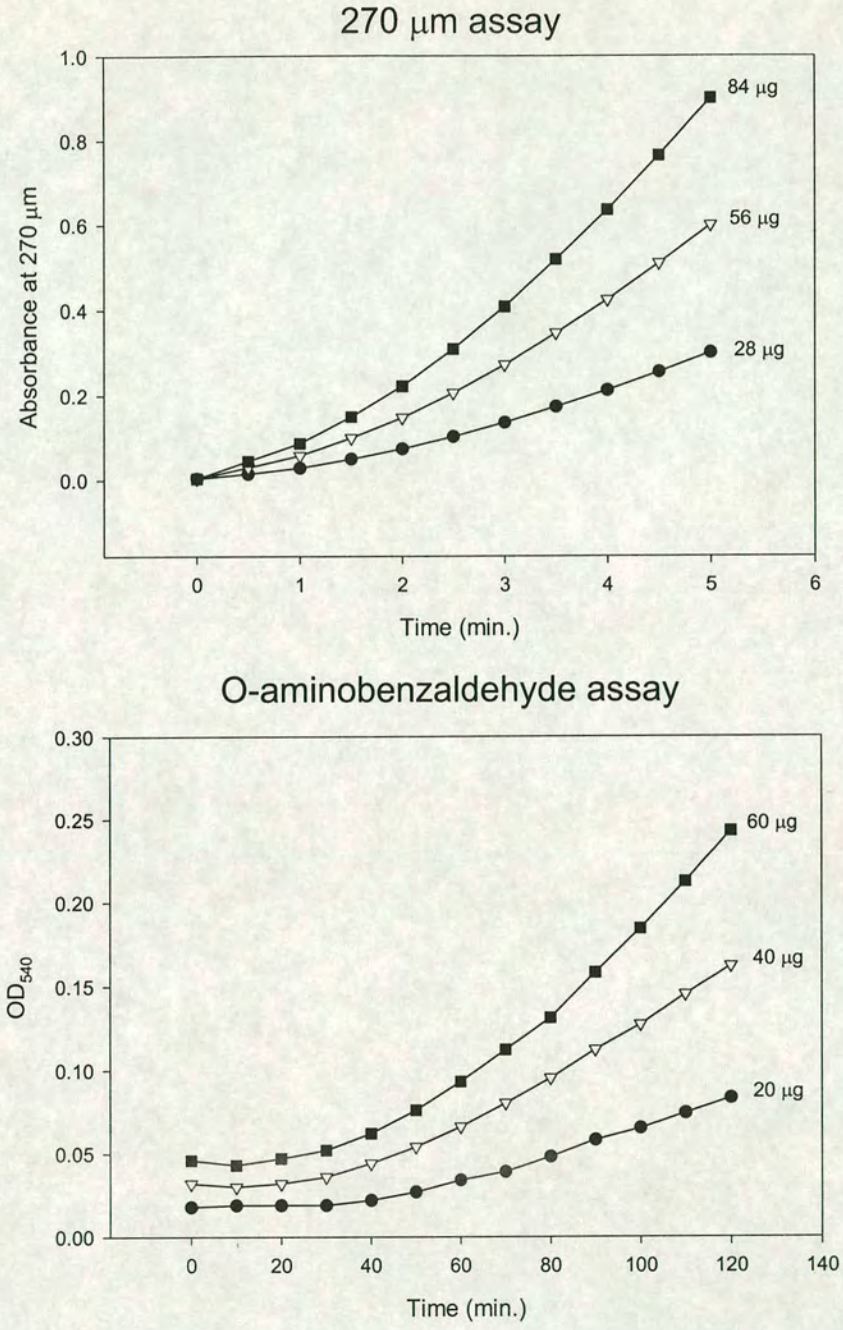


Fig. 2.5.1 Standard curves for both the 270 μm and o-aminobenzaldehyde assays. The amount of protein used is indicated next to each experimental curve. In both assays the gradient of the linear portion of the curve is directly proportional to the amount of protein added.

Chapter III

Effects of GroE depletion on DapA

3.1 Growth rates of MGM100

In the strain MGM100 the σ^{70} and σ^{32} promoters of GroE have been replaced with the *araC* gene and *araBAD* promoter from pBAD-18 (McLennan et al, 1998). This effectively puts GroE synthesis under the control of arabinose. Upon the addition of glucose to the growth medium, GroE production ceases and GroE levels halve with each generation.

To measure the growth rates of MGM100 the following protocol was used: MGM100 was grown overnight in 5ml LB + 0.2% ara and Kan. The following day the overnight culture was diluted 1/50 into 10ml LB + 0.2% ara and the culture was grown at 37°C with shaking until the OD_{600} reached ~ 0.2. 1ml of the culture was then removed, placed into an Eppendorf tube and centrifuged to pellet the cells. The supernatant was removed and the cells resuspended in either 10ml prewarmed LB + 0.2% ara or 0.2% glu. The cells were then returned to the 37°C water bath and the first sample (T=0) was taken. The cells were grown with constant agitation, taking samples as necessary (for a standard growth curve samples were taken every 30 minutes). If the OD_{600} of the cells went above 0.2 the cells were diluted 1/10 into fresh prewarmed medium.

Figure 3.1 shows typical growth curves of MGM100 grown under different conditions. The sample with arabinose present has a normal cellular complement of GroE and hence will continue to grow exponentially as long as the OD_{600} is kept below 0.2. The sample with glucose present has less GroE the longer it grows. For approximately the first 1.5 hours growth remains exponential (by the end of 2 hours, GroE levels will be reduced to 2-3% of normal levels), and then continues at a decreasing rate for another hour (at this time GroE levels would be reduced to ~ 0.25% of normal) before abruptly lysing, due to the absence of DAP in the peptidoglycan cell wall (McLennan et al, 1998). In the glucose sample + DAP, the growth rate is the same as the glucose sample, except the cells do not lyse at 2.5 hours. This is because exogenous DAP is being incorporated into the peptidoglycan, preventing cell lysis.

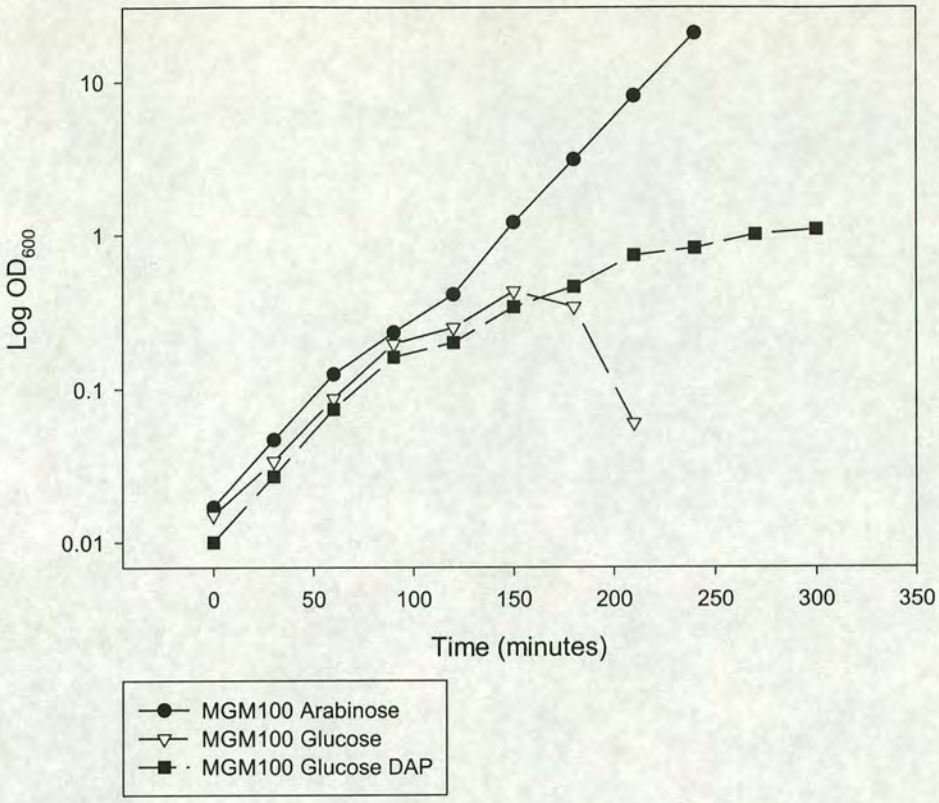


Figure 3.1 Growth rates of MGM100 in LB + ara ($GroE^+$), glu ($GroE^-$) or glu DAP ($GroE^-$, no lysis). For clarity, when cells have been diluted 1/10, the OD has been multiplied by 10 and plotted on the graph. 1/10 dilutions were performed at 60 minutes for all samples, at 120 minutes for the arabinose and glucose DAP samples and at 180 minutes for the arabinose sample.

3.2 DapA activity in GroE depleted cells

Previous work by McLennan (McLennan et al, 1998) suggested that DapA was affected by GroE depletion, causing GroE depleted cells to lyse. Therefore initial experiments were to determine if DapA activity was affected in GroE depleted cells. MGM100 was grown using the method given in 3.1 and at various times 5ml of cells (OD_{600} = 0.1) were removed, the cells were centrifuged at 1500 x g and the supernatant removed. The cells were resuspended in 250 µl of water and lysed by freeze-thaw lysis (Lenstra et al, 1983). The suspension was then centrifuged briefly to pellet any insoluble material and 100 µl of the cell extract was used in the o-aminobenzaldehyde assay, as described in materials and methods. A further 50 µl of the cell suspension was used to determine the protein concentration of the solution,

again using the protocol described in materials and methods, so that DapA activity could be normalised for all samples.

Figures 3.2.1, 3.2.2 and 3.2.3 show typical results from o-aminobenzaldehyde assays on protein samples removed at different times from cells grown in arabinose ($GroE^+$), glucose ($GroE^-$) and glucose plus DAP ($GroE^-$, no lysis) respectively. In the $GroE^+$ sample the lines of the graph are, on the whole, parallel to each other (any deviations in gradient are due to more or less sample being loaded, and when the gradient/protein concentration is determined all samples have the same activity per unit protein). Since the gradient of the curve (after the initial lag period) is proportional to DapA activity, this shows that in $GroE^+$ MGM100, DapA activity is constant. But in the $GroE^-$ samples, in the presence or absence of DAP, the later the samples are taken, the less DapA activity is recorded. This suggests that in MGM100, as the cells GroE deplete, DapA activity decreases, also the presence of DAP does not affect DapA activity. A tabulated form of these results is shown in Table 3.2.1

By normalising the gradients of the graphs against the protein concentrations of the respective samples, it is possible to plot the activity of DapA under various conditions. In figure 3.2.4 the average (gradient/protein concentration) was determined for all time points and then set as 100%. Using this it is possible to plot percentage DapA activity (relative to a $GroE^+$ sample) against the time the sample was taken. It is quite clear that in the $GroE^+$ sample DapA activity is constant, whereas DapA activity decreases markedly in GroE depleted cells, falling to ~10% of wild type just before lysis in the glucose sample.

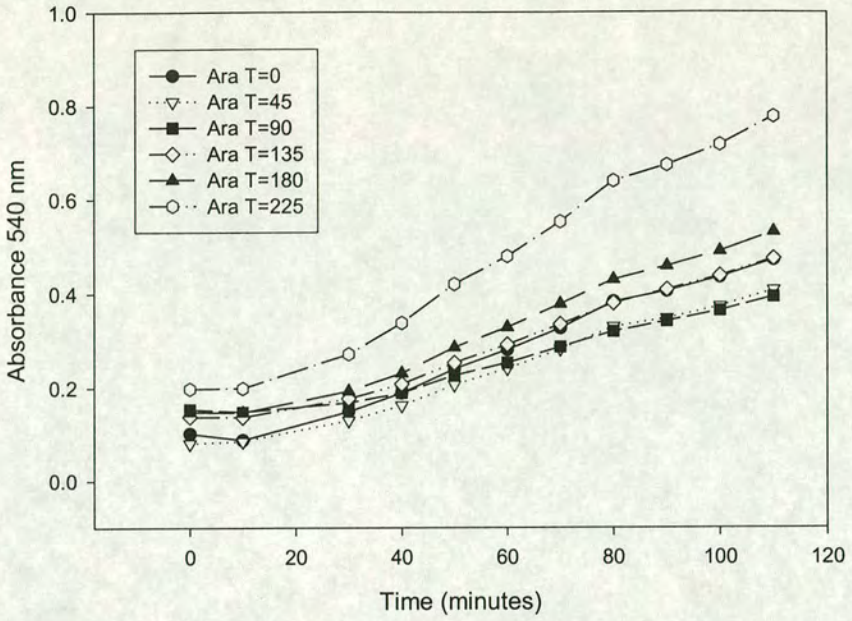


Fig. 3.2.1 Results of o-aminobenzaldehyde assay of DapA from MGM100 + arabinose ($GroE^{+}$).

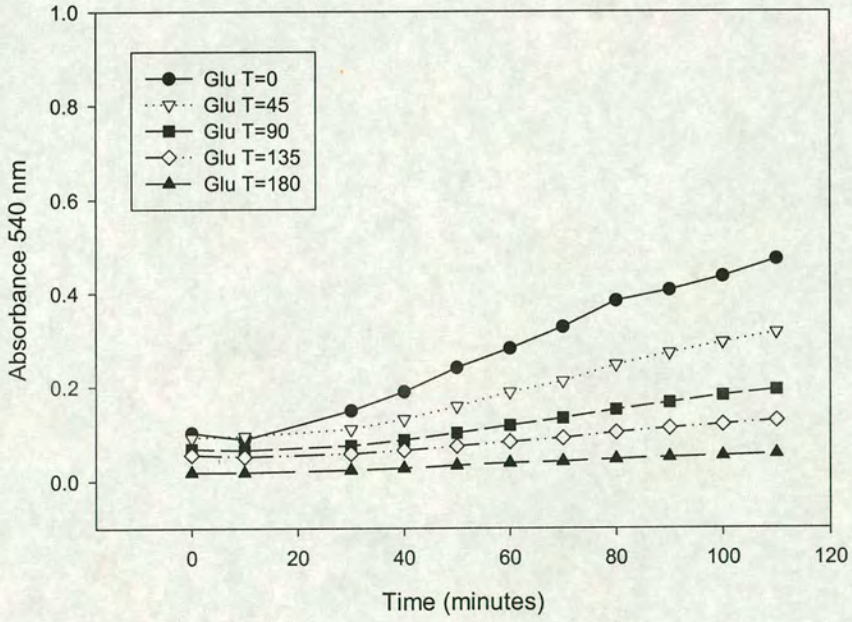


Fig. 3.2.2 Results of o-aminobenzaldehyde assay of DapA from MGM100 + glucose ($GroE^{-}$).

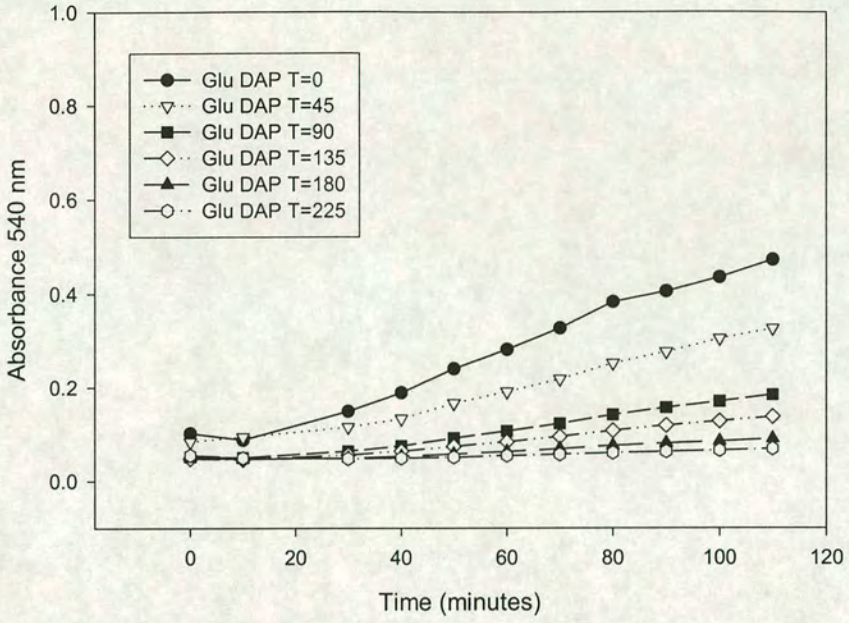


Fig. 3.2.3 Results of o-aminobenzaldehyde assay of DapA from MGM100 + glucose + DAP (GroE$^-$, no lysis).

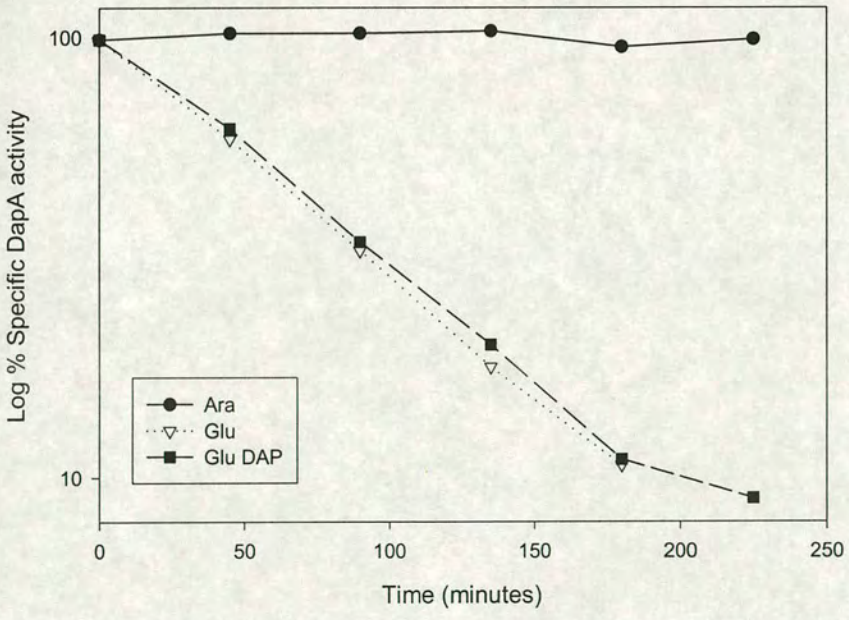


Fig. 3.2.4 DapA specific activity from MGM100 plotted against time of sampling. Samples taken from MGM100 grown in Ara (GroE$^+$), Glu (GroE$^-$) and GLU, DAP (GroE$^-$, no lysis).

Sample	Activity (gradient of slope)	Protein concentration of sample (mg ml^{-1})	Activity per mg of protein sample	Percentage activity [100% = average of Ara0 to Ara 225 (0.592)]
Ara 0	0.473	0.81	0.584	98.7%
Ara 45	0.435	0.72	0.604	102%
Ara 90	0.435	0.72	0.604	102%
Ara 135	0.489	0.8	0.611	103.2%
Ara 180	0.5	0.89	0.562	94.9%
Ara 225	0.79	1.35	0.585	98.8%
Glu 0	0.473	0.81	0.584	98.7%
Glu 45	0.286	0.825	0.347	58.6%
Glu 90	0.167	0.865	0.193	32.6%
Glu 135	0.1	0.95	0.105	17.7%
Glu 180	0.05	0.8	0.063	10.6%
Glu DAP 0	0.473	0.81	0.584	98.7%
Glu DAP 45	0.29	0.79	0.367	62%
Glu DAP 90	0.167	0.825	0.202	34.1%
Glu DAP 135	0.113	0.96	0.118	19.9%
Glu DAP 180	0.05	0.77	0.065	10.9%
Glu DAP 225	.055	1.04	0.053	8.9%

Table 3.2.1 DapA activity in MGM100 growing under various conditions: LB + Arabinose (Ara), LB + Glucose (Glu) and LB + Glucose DAP (Glu DAP). The numbers next to these letters represents the time in minutes at which the sample was taken. The activity of the sample is determined by the gradient of the linear portion of the graphs shown in figures 3.2.1-3.2.3. The protein concentration of the sample was determined, and the activity per mg of protein calculated. The percentage activity was determined with 100% being set at 0.592 (the average activity per mg of protein for all the arabinose samples).

3.3 DapA is an in vitro GroE substrate

The results in the previous section show that DapA activity is affected by GroE depletion. This result could be explained if DapA is a protein that requires GroE to reach its native conformation. Before determining if DapA interacts with GroE *in vivo* it was decided to perform refolding experiments of DapA *in vitro*.

0.3 μm DapA was denatured in the presence of guanidinium thiocyanate (GdmT) buffer [6.0M GdmT, 0.1M Tris-Cl pH 7.4 and 20mM dithioerythritol (DTE)] for 1.5 hours at 25°C (Buchner et al, 1991). This was found to denature DapA completely, so that it exhibited no activity when removed from the denaturant. The denatured protein was then diluted 100 fold into a refolding buffer (0.1M Tris-Cl pH 7.4, 10mM $MgCl_2$, 10mM KCl and 10 mM DTE, 2mM ATP) in the presence or absence of GroELS (2.5 μM) at 37°C. GroEL is a weak, potassium dependant ATPase (Viitanen et al, 1990), hence the presence of KCl and ATP in the refolding buffer.

After the denatured protein had been added to the refolding buffer, samples were taken out at various times and assayed for DapA activity. The samples were assayed using the o-aminobenzaldehyde assay in materials and methods, with one modification: 2 units of hexokinase and 10mM glucose were added to the assay mix. Hexokinase catalyses the transfer of a phosphoryl group from ATP to glucose to form glucose-6-phosphate (Anderson et al, 1975), thereby quenching ATP in the assay mix and preventing GroELS from folding DapA during the assay reaction.

The results of the refolding assay can be seen in figure 3.3.1. Undenatured DapA was used as a positive control, and for a negative control, denatured DapA was put into refolding buffer without GroELS. The graph shows quite clearly that DapA does not refold spontaneously to an active form when denatured, but can refold to an active form when GroELS (Sigma Aldrich) is present. Also we can determine from the graph that most DapA is refolded within the first 15 minutes of incubation with GroELS and then with a decreasing rate as the unfolded form becomes more dilute. Even if DapA is refolded in the presence of a greater concentration of GroELS the shape of the graph does not alter, this indicates that DapA and not GroELS is the limiting factor in this reaction. Once ~ 50% of DapA has been refolded to an active form the rate of refolding decreases dramatically, this may be because denatured DapA is now only 50% of the DapA population and therefore the possibility that it will interact with GroELS decreases.

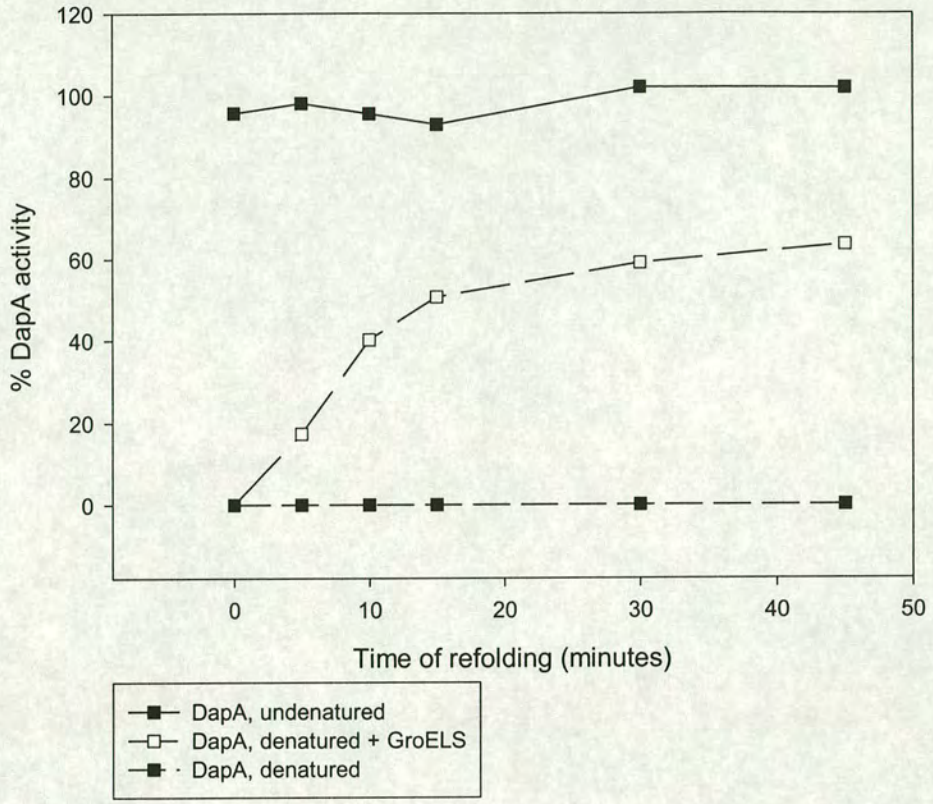


Fig. 3.3.1 Activity of denatured DapA that has been incubated with GroELS + ATP for various amounts of time.

This experiment was also repeated using only GroEL or GroES for refolding denatured DapA, but in both cases, no DapA activity was determined at any time point, suggesting that DapA requires both GroEL and ES for refolding *in vitro*.

3.3.1 Delayed lysis in cells over-expressing GroEL

Denatured DapA requires both GroEL and GroES to fold *in vitro* to a functional form; to determine if DapA also required GroELS to function *in vivo* MGM100 was transformed with a plasmid so that it could be depleted only of GroES. The plasmid pGroEL (Goloubinoff et al, 1989) contains a frame shift (inactivated) *groES* gene followed by a wild-type *groEL* gene cloned under the control of a P_{lac} promoter. Even in the absence of IPTG this plasmid causes the intracellular level of GroEL to increase 10 fold. The plasmid was transformed into MGM100, so that when the cells were grown in the presence of glucose, the chromosomal production of GroELS was repressed (being under the control of the promoter P_{BAD}) but the plasmid-based production of GroEL was unaffected (i.e. the cells were only depleted of GroES).

MGM100 pGroEL was grown, as per the protocol in section 3.1; chloramphenicol was also present to select for cells containing the plasmid pGroEL. The results of the growth curve are given in figure 3.3.2.

These results show that an increased level of GroEL only prevents MGM100 from lysing. Since lysis in GroELS depleted cells is caused by a deficiency in DapA activity, and since cells depleted of GroES alone do not lyse, it is possible that DapA is functional in GroES depleted cells. This then would suggest that DapA only requires GroEL to fold correctly *in vivo*.

To verify that pGroEL was not producing any GroES, the plasmid was transformed into the strain AG31 ($groES^{ts}$). The plasmid did not complement the mutation at the non-permissive temperature of 43°C.

Finally, DapA activity was measured in MGM100 depleted of either GroELS or GroES. The results (Table 3.3.1) show that while DapA activity does decrease in GroES depleted cells, it is not at the same rate as in GroELS depleted cells. The presence of more active DapA in GroES depleted cells may therefore be enough to prevent the cells from lysing.

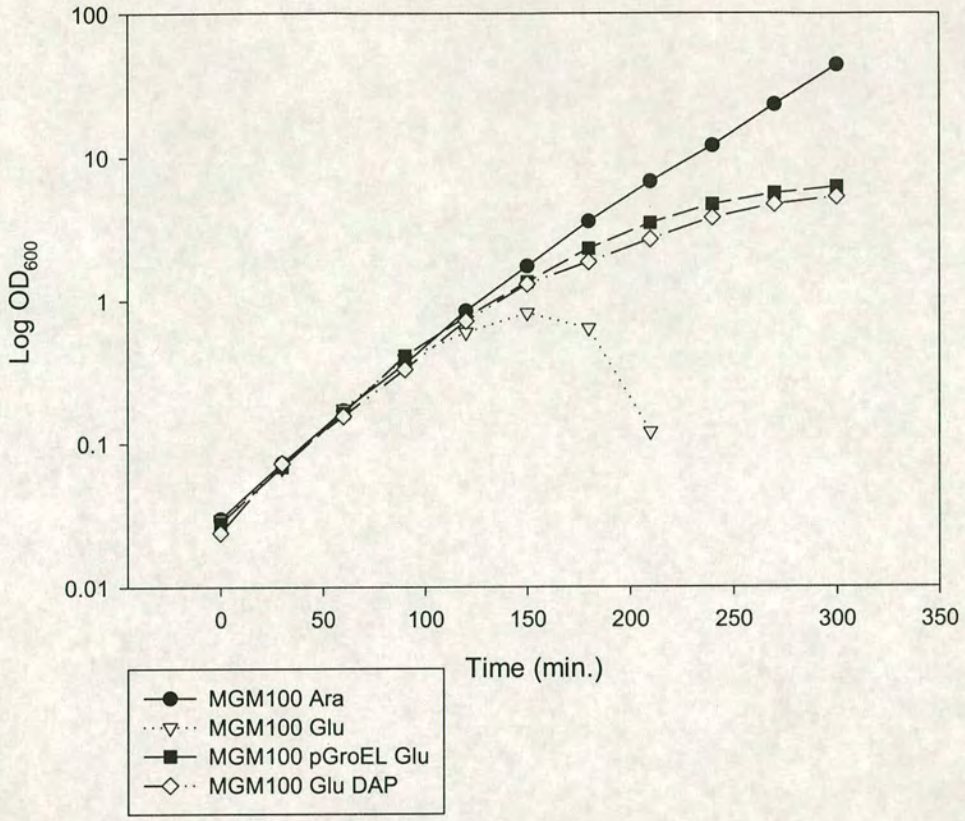


Fig. 3.3.2. Growth rates of MGM100 in LB + ara ($GroE^+$), glu ($GroE^-$) or glu DAP ($GroE^-$, no lysis). Also MGM100 pGroEL in LB + glu (GroES-). For clarity, when cells have been diluted 1/10, the OD has been multiplied by 10 and plotted on the graph. 1/10 dilutions were performed at 60 minutes for all samples, at 120 and 180 minutes for the arabinose, glucose DAP and pGroEL samples and at 240 minutes for the arabinose sample.

	Time after addition of glucose (hrs)				
	0	1	2	3	4
DapA activity in GroELS depleted cells (MGM100)	100	50%	24%	11%	-
DapA activity in GroES depleted cells (MGM100 pGroEL)	100%	85%	68%	53%	46%

Table 3.3.1 DapA activity in MGM100 or MGM100 pGroEL at various times after the addition of glucose (i.e. start of GroE depletion).

3.4 Pulse-chase experiments on DapA in MGM100

In previous studies of DapA by McLennan (McLennan et al, 1998) it was suggested that during GroE depletion in MGM100 the amount of DapA in the cell decreases. Since *dapA* promoter activity does not decrease in GroE depleted cells (N. McLennan, pers. comm., also this work, chapter 5) it was suggested that the turnover rate of newly synthesised DapA in GroE depleted cells must increase, relative to that in a $GroE^+$ cell.

To determine if this were the case MGM100 was grown as per the method in chapter 3.1 except the cells were grown in Spizizen's minimal medium instead of LB. Samples were taken from both a culture with arabinose ($GroE^+$) and a culture with glucose and DAP ($GroE^-$), the $GroE^-$ sample being taken at a time such that the intracellular level of GroE was less than 1% of the original cellular complement. 10-$\mu Ci\ ml^{-1}$ of ^{35}S-methionine was then added to the samples and they were incubated at 37°C for 5 minutes. During this time newly synthesised proteins would incorporate the radioactive methionine and would therefore become radioactively labelled, this would include DapA, which contains nine methionine residues. After the five minute labelling period a solution of non-radioactive methionine was added such that it was in great molar excess ($\sim 10^8$) compared to the radioactive methionine. This chase prevents any further labelling of the proteins with radioactive methionine.

By taking samples at various times after the addition of the non-radioactive methionine and examining them by 2-D electrophoresis (following the protocol in materials and methods) it is possible to determine the turnover rate of proteins of interest. The location of DapA on a 2-D gel of *E. coli* proteins could be determined using the data provided by the Swiss-2D database (http://www.expasy.ch). If DapA did have an increased rate of turnover in GroE depleted cells, then the spot on the 2-D gel corresponding to DapA should disappear faster in a GroE depleted sample that in a $GroE^+$ sample.

The results of the 2-D analysis are shown in figure 3.4.1. The gels were normalised by comparing the intensity of 5 spots (thereby reducing the possibility of error introduced by only comparing one spot), then the intensity of the DapA spot was determined at 0 minutes and 20 minutes after the addition of non-radioactive methionine. In a $GroE^+$ cell, after 20 minutes, there is still ~ 88% of the original DapA still present, while in a $GroE^-$ cell, it is only slightly lower at ~ 83%. It can be concluded that DapA does not appear to have an increased level of turnover in GroE depleted cells.

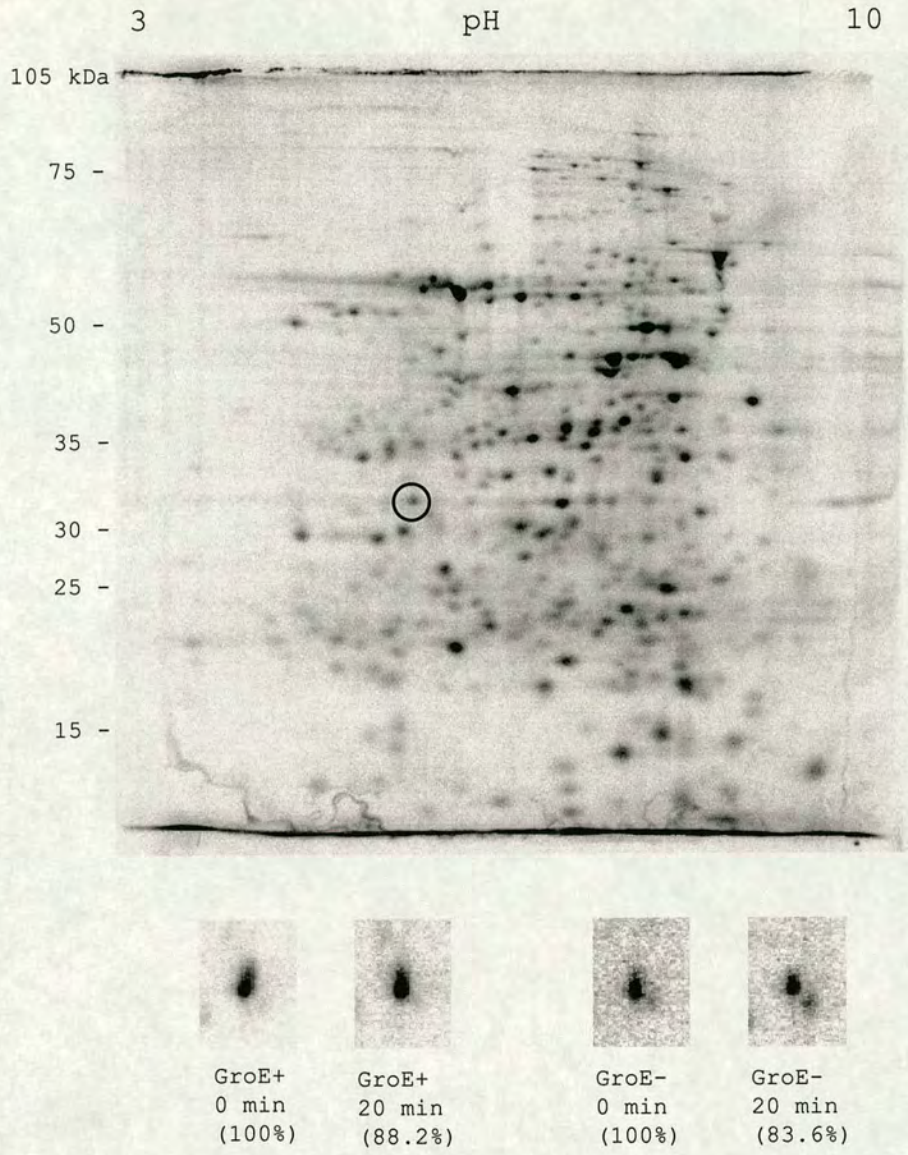


Fig. 3.4.1 2-D gel of proteins extracted from *E. coli*. In the upper picture the spot representing the protein DapA is circled. In the lower set of pictures is the DapA spot from cells grown in the presence of arabinose ($GroE^+$) or glucose ($GroE^-$), which have been incubated with 10 μCi ml^{-1} of ^{35}S-methionine for 5 minutes, then had an excess of non-radioactive methionine added. Protein samples were taken at the stated time after the addition of the non-radioactive methionine and the intensity of the DapA spot was determined.

3.5 DapA turnover is not increased in GroE depleted cells

The pulse-chase experiments with DapA had the disadvantage that the cells were growing in Spizizen's minimal medium rather than LB; the medium in which Dr. McLennan *et al* observed a decreased level of DapA in GroE depleted cells. To verify if the results observed in the pulse chase experiment were correct, MGM100 was again grown in LB, using the protocol given chapter 3.1 and samples were taken at different stages of growth from both $GroE^+$ and $GroE^-$ cells. These samples were analysed by Western blotting followed by detection with anti-DapA antibodies (a gift from Dr. Galili, Weizmann Institute of Science) and a subsequent chemiluminescent step.

In the previous work by McLennan *et al.* to correct for possible loading differences in the SDS-PAGE step of the procedure, the level of FtsZ was also measured in the samples, using anti-FtsZ antibodies. It was assumed that FtsZ levels were not affected by GroE depletion, and therefore by normalising each sample so that it contained the same amount of FtsZ, DapA levels could be accurately measured.

When the Western blotting procedure was repeated in this work, the protein concentrations of the samples were determined and appropriate amounts were loaded onto the SDS-PAGE gel accordingly. When the detection procedure was complete it was found that DapA levels did not decrease in GroE depleted cell, rather the levels of FtsZ increased as the cells became more GroE depleted (Fig. 3.5.1). Further evidence for this observation is that in MGM100 grown in LB + glucose, in the absence of DAP, DapA levels are actually increased just before lysis, which agrees with the observation that *dapA* promoter activity increases just before lysis in this sample (chapter 5).

For the analysis of FtsZ and DapA levels by Western blotting it was important to expose the samples long enough for them to be visible on X-ray film, but not long enough to saturate the film (i.e. working outside the linear range of the X-ray film). To overcome this problem, serial dilutions of the samples were used and also the samples were exposed to X-ray film for varying amounts of time. The films were then analysed by densitometry. Details of a typical set of results can be seen in Table 3.5.1.

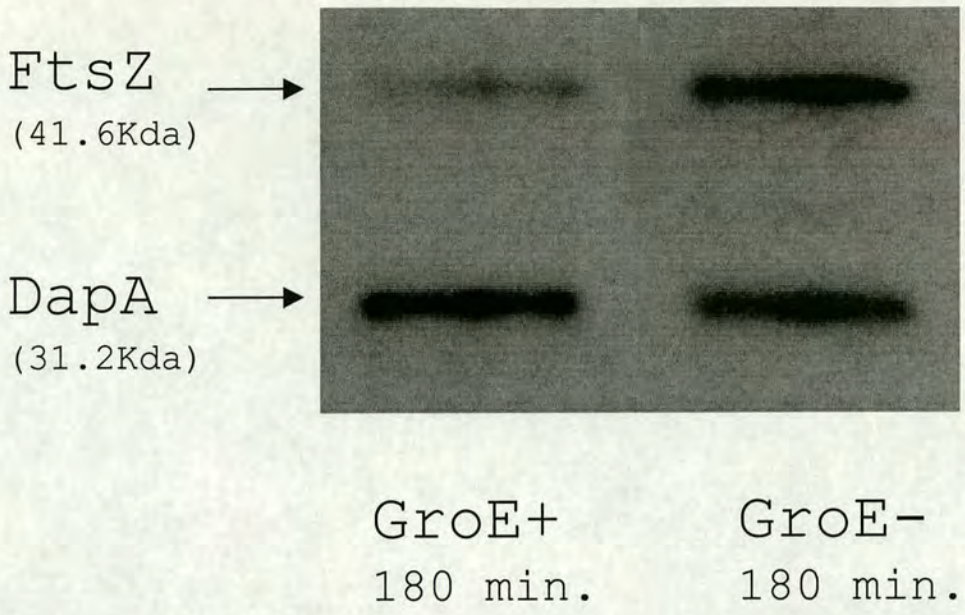


Fig. 3.5.1 Western blot of DapA from MGM100 grown in LB + arabinose ($GroE^+$) or LB + glucose, DAP ($GroE^-$) for 3 hours. Total protein concentration was determined for each sample, and equivalent amounts loaded. The level of the protein DapA remains constant independent of cellular GroE levels, while levels of FtsZ increase in GroE depleted cells. The $GroE^-$ sample has 2.3 x more FtsZ than the $GroE^+$ sample, relative to DapA.

FtsZ transcription is under the control of a σ^S promoter, *ftsQ1p* (Vicente et al, 1998). The sigma factor σ^S, encoded by the gene *rpoS*, is switched on during the stationary phase of growth in *E. coli*, such that *rpoS* is transcribed at a level 20 times higher than in logarithmic growth (Lowen et al, 1994). In MGM100 growth rate decreases as the cells become more GroE depleted (Fig. 3.1), this could plausibly cause an increased expression of *rpoS* and a corresponding increase in expression of *ftsZ*. Therefore by normalising FtsZ levels in GroE depleted cells, a decreased DapA level will be observed, but this is not actually representative of DapA levels in the cell, which, in fact, remains constant.

After analysis of 44 different exposures (from three different experiments varying the time of exposure, and also using the results from a number of serial dilutions) it was determined that after 3 hours of GroE depletion, MGM100 growing in LB had 1.3 ± 0.2 times more DapA and 4.08 ± 0.04 times more FtsZ than it did after 30 minutes of GroE depletion.

Sample time	30 min.		180 min.			
Sample dilution	1/4 (A)	1/8 (B)	1/4 (C)	1/8 (D)	(C)/(A)	(D)/(B)
DapA	1874.7	918.8	1933.6	1029	1.03	1.12
FtsZ	374.0	206.6	1825.2	971.2	4.88	4.70

Table 3.5.1 Results of a typical analysis of DapA and FtsZ levels by densitometry in GroE depleting cells. Samples were taken at an early (30 minutes) or late (180 minutes) time during GroE depletion. The samples were diluted appropriately so that exposure to X-ray film would not cause saturation. The DapA and FtsZ levels were determined and also the ratio of the late time point to the early time point. In this example there is ~ 1.08 times more DapA and ~ 4.8 time more FtsZ at 180 minutes than at 30 minutes in GroE depleted cells.

3.6 Gene replacement of *dapA*

Since it had been shown that DapA protein was still present in GroE depleted cells, it was necessary to determine why DapA activity decreased as the intracellular levels of GroE decreased. Native DapA exists as a homotetramer with a molecular weight of 31,372 per monomer (Shedlarsky et al, 1970; Richaud et al, 1986). Since DapA was being transcribed and translated, yet showed no enzymatic activity in GroE depleted cells, it was possible then that the enzyme was not folding to its correct native state. Since previous Western analysis of DapA in MGM100 had been under denaturing conditions, it was unknown whether DapA forms a homotetramer in GroE depleted cells. It was hypothesised that if DapA were a GroE substrate, then in the absence of GroE newly synthesised DapA may not be able to fold correctly, and therefore may not reach its correct homotetrameric native state.

To determine if DapA was indeed only present as either a monomeric species or in an aggregated form in GroE depleted cells, it was necessary to analyse the protein on a native gel. Unfortunately, the supply of antibodies that had been used previously to detect DapA had been exhausted, and it was not possible to obtain any more; therefore a new method to detect DapA had to be devised. A solution to this problem was to place a 'tag' onto the carboxy terminus of DapA, which would serve as an antigen for commercially available antibodies.

The theoretic construction of a tagged version of DapA would be to construct the 'tagged' *dapA* gene on a plasmid and then use this to replace the wild type *dapA* gene on the chromosome of MGM100. A number of gene replacement strategies exist but in this case a novel method developed by N. McLennan and M. Masters was employed. This gene replacement relies on the fact that high copy-number ColE1 based plasmids, such as the pUC range of vectors, appear to recombine with the host chromosome via a single crossover and excise by the same mechanism. Regions of homology between the plasmid and the chromosome are required for the crossover to occur. The excision of the plasmid appears to be promoted by instability caused by the presence of a plasmid origin of replication on the *E. coli* chromosome.

It is possible to direct the location of insertion of the plasmid onto the chromosome and interrupt or replace the gene of interest by cloning the gene, which has been inactivated by the insertion of a selectable marker such as an antibiotic resistance cassette, into the high copy-number plasmid. It is proposed that the plasmid can insert into the locus of interest via a single crossover between homologous DNA. The amount of flanking DNA required for efficient integration has not yet been determined, although 200 bp of homologous DNA upstream and downstream of the resistance marker has been shown to suffice (P. Stumpfle, pers. Comm.). Upon excision, one of two plasmids can be formed: a molecule identical to the plasmid that was initially introduced into the cells or a plasmid containing the wild-type locus, with the chromosomal ORF now interrupted by the resistance marker (Fig. 3.6.1).

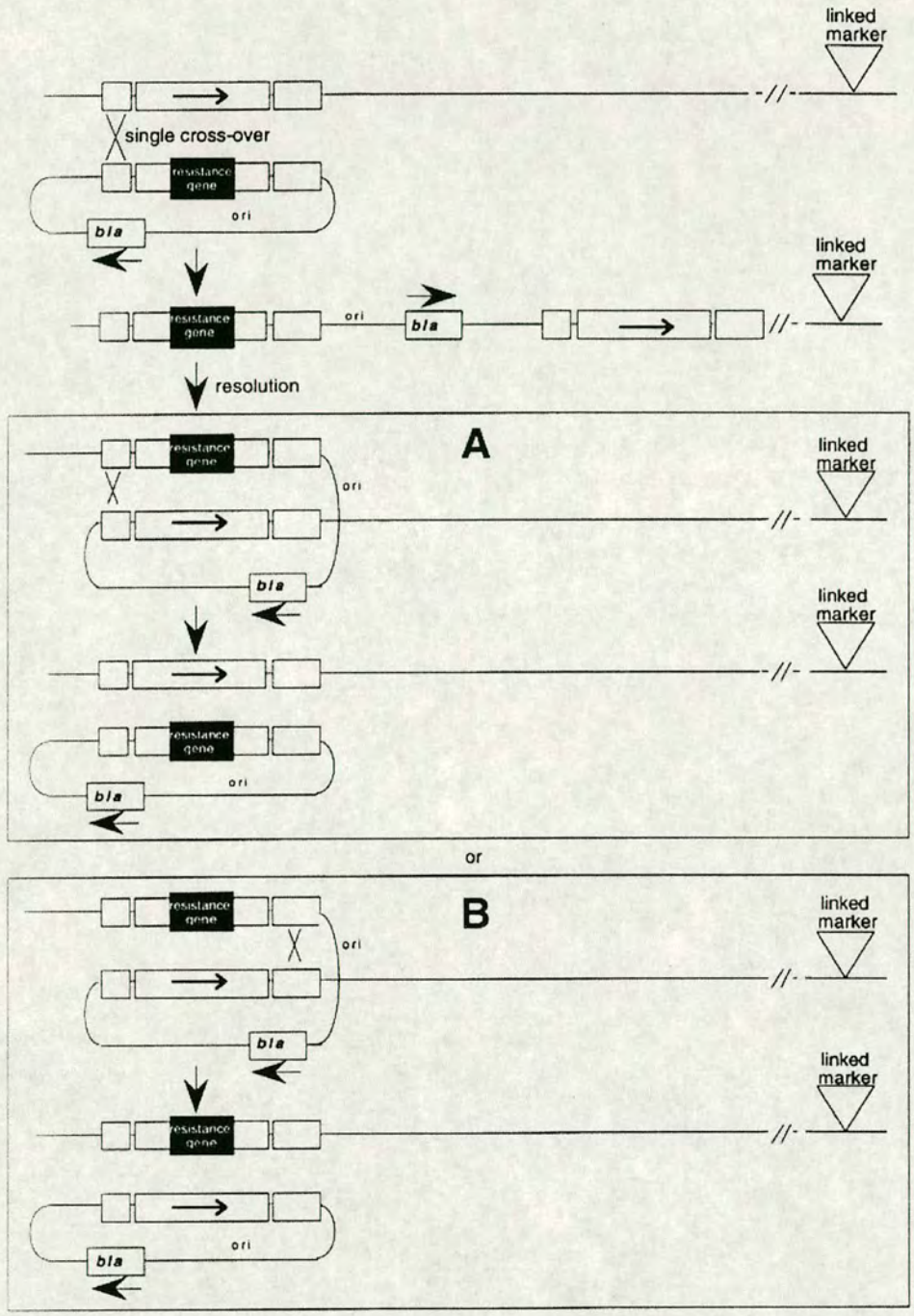


Fig. 3.6.1 The single cross-over of plasmid DNA at a homologous region of chromosomal DNA. Excision of the plasmid can result either in the reformation of the original plasmid or the antibiotic marker can remain on the chromosome and the wild-type locus is transferred onto the plasmid.

A P1 lysate is grown on the strain carrying the plasmid to be used for the replacement, the strain should also have an additional selectable marker within transducing range (<2 minutes) of the chromosomal gene being replaced. The transducing particles containing the region of interest will carry three possible arrangements of the chromosome: (i) the wild type locus, (ii) the locus with the entire plasmid inserted or (iii) the locus with the ORF of choice interrupted by the antibiotic marker.

The P1 lysate is then used to transduce a wild-type strain, selecting for the insertion marker and the linked chromosomal marker. The resultant colonies must then be screened to ensure loss of the original plasmid. This is achieved by screening for the loss of the plasmid resistance marker, in the case of the pUC vectors, the loss of Amp resistance.

3.6.1 Construction of plasmid pUCDap·HSV

The HSV·Tag (Novagen) is a 36mer, which encodes a peptide of 11 amino acids derived from Herpes Simplex Virus glycoprotein D:

```
   GlnProGluLeuAlaProGluAspProGluAspEND
5′ CAGCCTGAACTCGCTCCAGAGGATCCGGAAGATTAA 3′
3′ GTCGGACTTGAGCGAGGTCTCCTAGGCCTTCTAATT 5′
```

There is a commercially available HSV·Tag mouse monoclonal IgG_1, that has high specificity and affinity for the peptide encoded by this sequence (Novagen). To determine if the HSV·Tag was suitable for use with MGM100, a whole cell extract of MGM100 and a positive control extract containing a 31,102 Dalton protein that carries a C-terminal HSV·Tag was Western blotted and probed with anti-HSV antibody, as per the protocol described in materials and methods. The positive control gave one clear band, and there were no bands present in the MGM100 cell extract lane, indicating that the HSV antibody has no cross-reactivity with any endogenous MGM100 proteins, and was therefore suitable for use.

The HSV tagging of *dapA* is a three-step process. Firstly a pUC18 plasmid containing a 4 kb fragment of DNA (pUCDapA) containing the coding sequence of *dapA* and at least 1 kb of the upstream and downstream genes *nlpB* and *gcvR*, was subjected to a site directed mutagenesis. The purpose of this was to replace the *dapA*

stop codon TAA with a restriction site, so that the HSV·Tag sequence could be ligated in frame to produce the fusion gene *dapA·HSV*.

pUCDapA was analysed and found to contain no *Bgl*II restriction sites. Two primers, DAPBGLTOP and DAPBGLBOT, were designed which, when used with the site directed mutagenesis protocol described in materials and methods would mutate the *dapA* stop codon to a *Bgl*II restriction site.

DAPBGLTOP: 5′ATCAAATCTCCCTAAAGATCTCAGCAAACCGGCATG 3′
DAPBGLBOT: 5′CATGCCGGTTTGCTGAGATCTTTAGGGAGATTTGAT 3′

dapA terminal sequence before mutagenesis:

5′ AAGCATGCCGGTTTGCTG**TAA**AGTTTAGGGAGATTTGATGGCT 3′
dapA..LysHisAlaGlyLeuLeuEND MetAla..*nlpB*

dapA terminal sequence after mutagenesis:

5′ AAGCATGCCGGTTTGCTGAGATCTTTAGGGAGATTTGATGGCT 3′
*Bgl*II

When the point mutation was completed the plasmid pUCDapA, which previously had no *Bgl*II restriction sites was now linearised when cut with the restriction enzyme *Bgl*II (Fig. 3.6.2). The mutated plasmid was now called pUCDapA(mut).

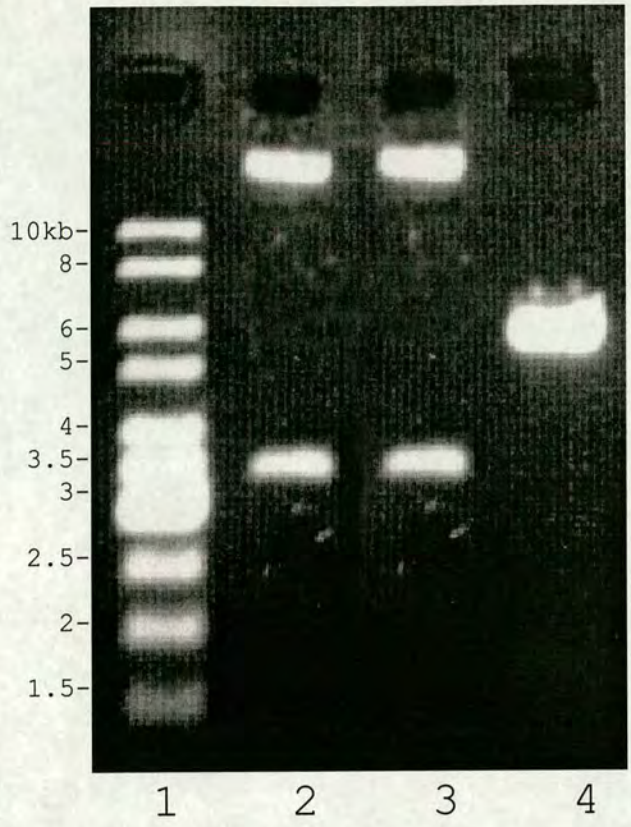


Fig. 3.6.2 Restriction analysis of pUCDapA(mut). Lane 1; DNA markers, lane 2; pUCDapA, lane 3; pUCDapA incubated with *Bgl*II, lane 4; pUCDapA(mut) incubated with *Bgl*II. The plasmid pUCDapA has no *Bgl*II restriction sites, but after point mutation has 1 unique *Bgl*II site replacing the TAA stop codon of *dapA*. The bands in lanes 2 and 3 represent the supercoiled (small band) and open circle (large band) form of pUCDapA.

The second stage in the tagging process was the construction of the HSV·Tag itself. In addition to the HSV·Tag being ligated in frame to *dapA*, an antibiotic resistance gene, *cat*, providing Cmp^R, would also be inserted next to the tagged gene, providing an easy screening method for the gene when it was placed onto the chromosome. This was achieved by making a PCR product that contained *Bgl*II sites on its 3′ and 5′ ends, the HSV·Tag sequence and the sequence of the *cat* gene (including 955 bp of upstream sequence so that it would be under the control of its native promoter).

The PCR required two primers, the first primer, HSVCATTOP contains a *Bgl*II restriction enzyme recognition site, the coding sequence for the HSV·Tag and a region of DNA complementary to the DNA sequence 955 bp upstream of the start site of transcription for the *cat* gene on pBR325.

HSVCATTOP:

5′.GAGTAAGTAGATCTCAGCCTGAACTCGCTCCAGAGGATCCGGAAGATTAA**GCA-**

*Bgl*II HSV·Tag END

-GGCATCGTGGTGTCACG.3′

Complementary

Region

The second primer HCATBOT annealed 372 bp downstream of the TAA stop codon of the *cat* gene in pBR325 and introduced a *Bgl*II restriction enzyme recognition site.

HCATBOT: 5′.AGGCGCCGCCCTAGATCTTGTCTGCCTCCC.3′

*Bgl*II

The two primers were used in a PCR with the plasmid pBR325 as the template DNA, according to the protocol given in materials and methods. The final product was 2064 bp long and was sub cloned into pGEM-T.

The pGEM-T Easy Vector systems (Promega) are convenient systems for the cloning of PCR products. The vectors are prepared by cutting pGEM-T with the restriction enzyme *Eco*RV and adding a terminal 3′ thymidine to both ends. These 3′-T overhangs greatly improve the efficiency of ligation of a PCR product into the plasmid by preventing recircularisation of the vector and providing a compatible overhang for PCR products generated by thermostable polymerases which add a single deoxyadenosine in a template independent fashion to the 3′-ends of the amplified fragments.

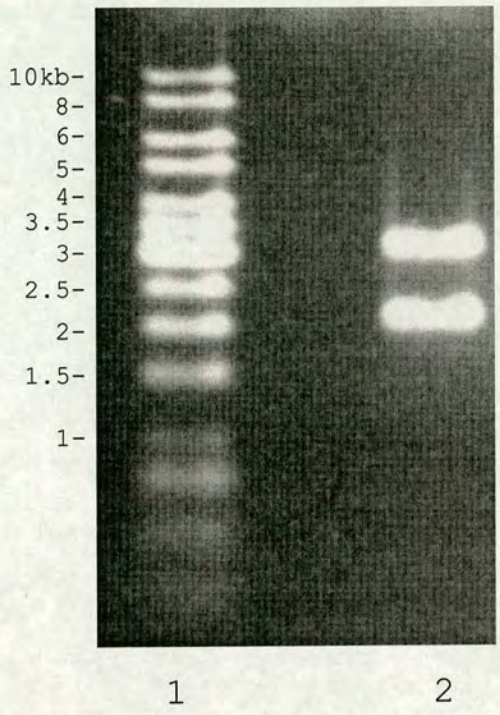
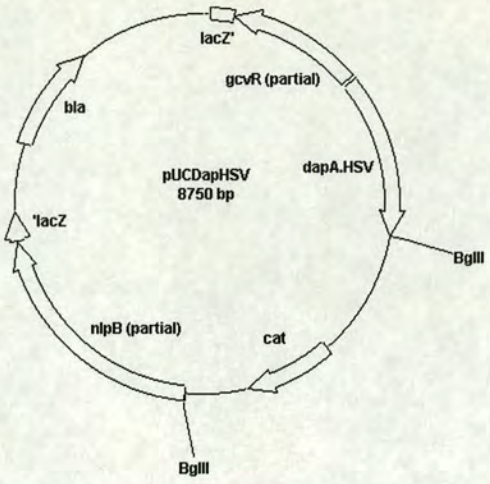


Fig. 3.6.3 Restriction of pGEM-T HSV*cat* with *Bgl*II. Lane 1: DNA markers, lane 2: pGEM-T HSV*cat* restricted with *Bgl*II to give 2 fragments, pGEM-T (3.0kb) and HSV*cat* (~ 2kb).

The pGEM-T plasmid containing the 2064 bp insert was transformed into TOP-10 and selected on plates containing Cmp and Amp and grown overnight at 37°C. Transformed colonies were picked and the plasmid DNA extracted according to the protocol in materials and methods, the plasmid was then restricted with the enzyme *Bgl*II and run on an agarose gel (Fig. 3.6.3). The ~ 2kb fragment, corresponding to the HSV-*cat* cassette was extracted from the agarose gel using the β-agarase protocol in material and methods, then the cleaned DNA was ligated to pUCDapA(mut) that had also been restricted with *Bgl*II and CIP treated, to prevent plasmid re-ligation. The ligation reaction was transformed into TOP-10 and transformants were selected on LB plates containing Amp and Cmp overnight at 37°C.



5′....AAGCATGCCGGTTTGCTGAGATCTCAGCCTGAACTCGCTCCAGAGGATCCG
LysHisAlaGlyLeuLeuArgSerGlnProGluLeuAlaProGluAspPro
dapA *BglII* HSV·Tag

GAAGATTAA**GCAGGCATCGTGGTGTCACG**....3′
GluAspEND
END Upstream *cat*

Fig. 3.6.4. The plasmid pUCDapHSV (details of construction in text). The lower text shows the sequence of the plasmid at HSV·Tag region and shows that, when translated, DapA will have a HSV·Tag.

As the plasmid pUCDapA(mut) was restricted with only one enzyme (*Bgl*II) it was possible that the HSV*cat* insert could ligate to the plasmid in either orientation. If the cassette ligated in the correct orientation (Fig. 3.6.4) then the HSV·Tag would be in frame with *dapA* and form a fusion protein, but if the cassette ligated in the opposite orientation there would be no tagged *dapA*. To determine the orientation of the cassette two primers were designed, one of which (DHCHKDAP) would anneal to DNA sequence in the *dapA* gene, and the other (TAGCHK) would anneal to DNA sequence in the upstream region of the *cat* gene.

DHCHKDAP: 5′ GAAGGGCATTTTGCCGAGGC 3′
TAGCHK: 5′ TCCGTGTCGCCCTTATTCCC 3′

Only if the HSV*cat* cassette had ligated in the correct orientation would it be possible to obtain a PCR product of 781 bp. 10 colonies were checked using this method and six were found that contained a plasmid (pUCDapHSV) which gave the correct size of fragment when checked by PCR.

3.6.2 Construction of plasmid pUCDap·His

A 6 x His tagged form of *dapA* was also constructed, to aid in the purification of DapA from *E. coli*. The construction of this plasmid, pUCDapHIS, was essentially the same as for pUCDapHSV, except instead of using the primer HSVCATTOP (which introduced the HSV sequence), the primer HISCATTOP was used.

HISCATTOP:

5′...CAGTTAATAGATCTCATCACCATCACCATCACTAA**GCAGGCA**

*Bgl*II HisHisHisHisHisHisEND

TCGTGGTGTCACG...3′

Complementary

Region

Because of the slightly smaller length of the primer compared to HSVCATTOP, when the plasmids pUCDapHIS were being screened using the primers TAGCHK and DHCHKDAP, the product would be 766 bp if the HIS·Tag cassette were ligated in the correct orientation.

3.6.3 Construction of pUCDapA::CAT

To determine if the tagging of DapA interfered with the activity of the protein, it was necessary to construct a *dapA* delete strain of *E. coli*, and determine if the plasmids carrying the tagged DapA could complement this deletion.

The construction of the *dapA* deletion cassette used the crossover PCR technique (Link et al, 1997). Crossover PCR was originally used by Link *et al.* for the deletion of genes. Briefly, the protocol comprises three steps as illustrated in

figure 3.6.5. In the first step two different asymmetric PCRs were used to generate fragments to the left and right of the sequences targeted for deletion. The two inside primers, B and C, are complementary to the 5′ and 3′ ends of the gene to be deleted, but they also contain a 'tail' region, which is complementary over 21 nucleotides and may contain a restriction site. The outside primers, A and D were used to introduce restriction enzyme recognition sites into the flanking genes so that the entire final product could be easily cloned. Two PCR products were made using primer pairs AB and CD respectively, when these two PCR products are mixed the complementary regions anneal and prime at the 3′ overlapping region for a 3′ extension of the complementary strand. The fused molecule is then amplified using the outside primers A and D to give the final product with the gene deleted and restriction sites on either end. Since a restriction site was also introduced into the middle of the product, an antibiotic resistance gene could be inserted here for ease of selection when the deletion construct is moved onto the chromosome.

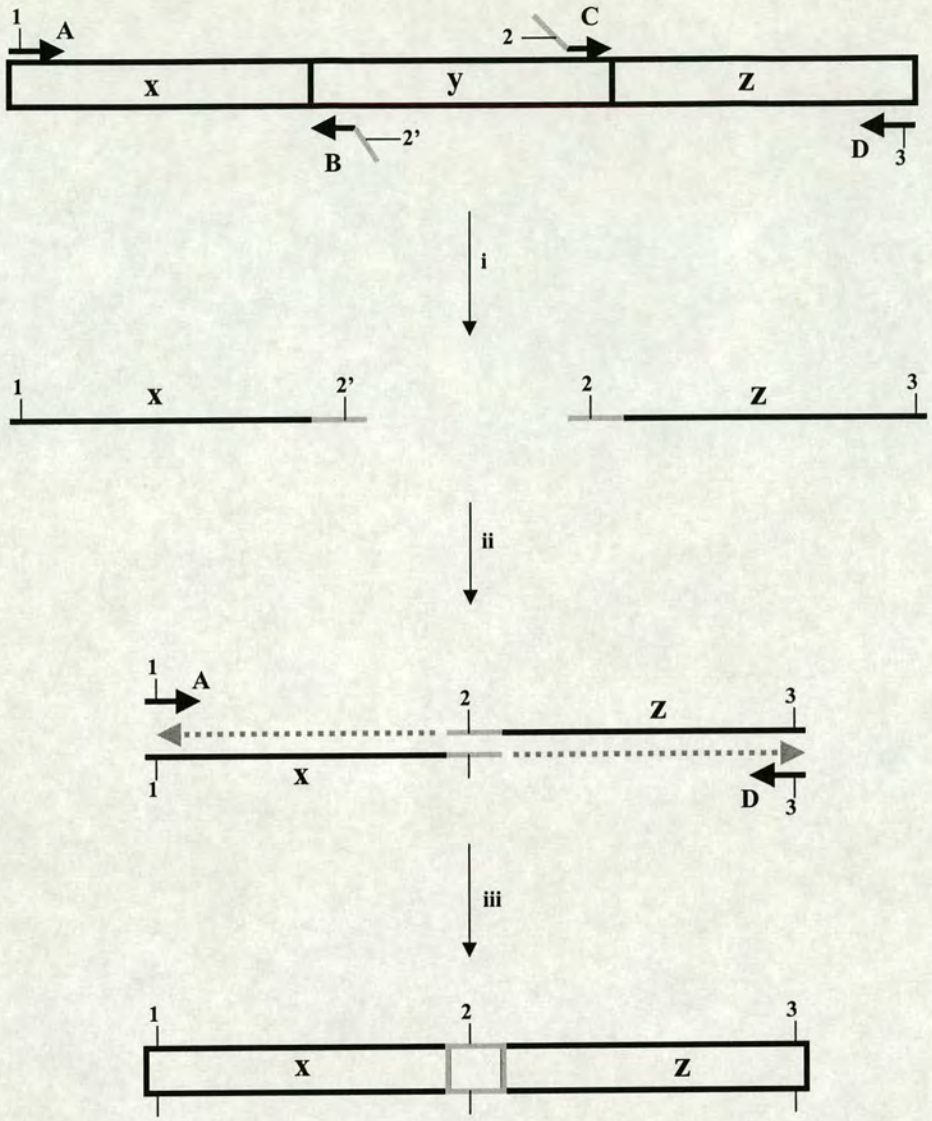


Fig. 3.6.5 Summary of cross-over PCR. x: upstream flanking gene, y: gene to be deleted, z: downstream flanking gene, A: outside primer , B: inside tailed primer, C: inside tailed primer, D: outside primer, 1: restriction site, 2: restriction site (not the same as either 1 or 3), 3: restriction site (may be the same as 1). i; PCRs of flanking genes using primer pair AB and CD to give products with complementary 3′ ends. ii: PCR products mixed, 3′ ends anneal and primer extension occurs, add outside primers A and D. iii: PCR with primers A and D to give final product with deleted gene y. Product can be restricted with enzymes 1 and 3 and ligated to an appropriate vector. Product may then be restricted with enzyme 2 and an antibiotic resistance cassette with complementary ends may be ligated in place.

To make the Δ*dapA* cassette four primers were designed, the two outside primers, End 5-3 and End 3-5 annealed to DNA located in the *nlpB* and *gcvR* genes respectively and introduced mutations so that the final PCR product would have a 5′ *Bam*HI and 3′ *Eco*RI restriction enzyme recognition site. The two inside primers END 5-3 and END 3-5 contained a region of homology to the 5′ and 3′ ends of the *dapA* gene and also the complementary 'tails' containing a *Sma*I/*Xma*I recognition site (the sequence of the 'tail' region was taken directly from the paper by Link *et al.* which first described this method).

END 5-3: 5′ AATCGCGGGATCCTGCAGACGCAC 3′
*Bam*HI

END 3-5: 5′ CGCTGTCCCTGGAATTCGGTCAAA 3′
*Eco*RI

DAP 5-3: 5′ **CCCATCCCACAAGCCCGGGACA**CCCGTGAACATGGGCCATCC 3′
Tail Region — *Sma*I/*Xma*I — Homology to *dapA* region

DAP 3-5: 5′ **TGTCCCGGGCTTGTGGATGGG**AAGCATGCCGGTTTGCTGTA 3′
*Sma*I/*Xma*I — Tail region — Homology to *dapA* region

The primers END5-3 and DAP3-5 were used in a PCR reaction, as were the primers END3-5 and DAP 5-3, using chromosomal DNA extracted from MG1655 as the template. The two products were 1170 bp and 844 bp respectively. The PCR products were run on an agarose gel, and the DNA was extracted using the β-agarase protocol as described in materials and methods. 10 ng of each product were then mixed together and used as the template in a new PCR using the primers END 5-3 and END 3-5, their complementary 'tail' regions would anneal to each other to produce the Δ*dapA* cassette, a product of 2014 bp (Fig. 3.6.6).

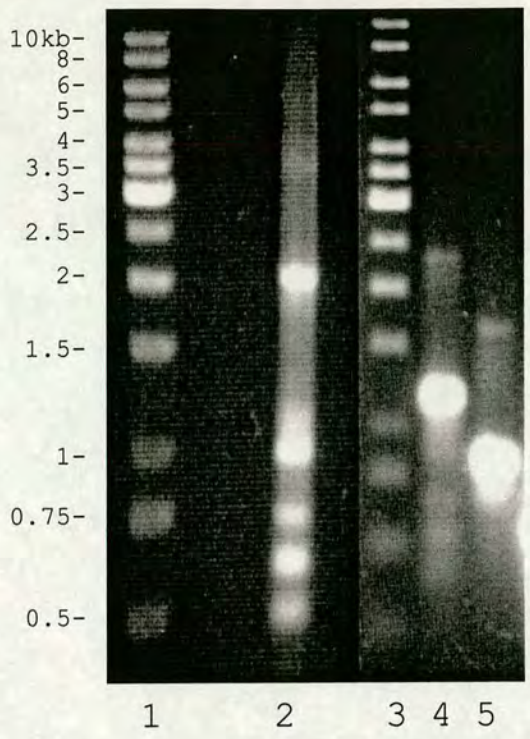


Fig. 3.6.6 Cross-over PCR products for the construction of Δ*dapA* cassette. Lanes 1 and 3: DNA markers, lane 2: cross-over PRC product (2014 bp), lane 4: product of END5-3/DAP3-5 PCR (1170 bp), lane 5: product of END3-5/DAP5-3 PCR (844 bp).

The Δ*dapA* cassette PCR product (2014 bp) was run on an agarose gel, extracted using the β-agarase method, and sub cloned in pGEM-T. The construct was transformed into TOP-10 competent cells and plated on LB + Amp + X-gal, and screened for blue or white colonies. Since cloning a cassette into pGEM-T disrupts the *lacZ* gene, resultant colonies will be unable to hydrolyse X-gal, and will therefore produce white colonies. A number of white colonies were picked and the plasmid DNA was extracted and subjected to restriction analysis to determine if the construct was correct. When a plasmid bearing the correct construct was found, the cassette was released from the plasmid by digesting with the restriction enzymes *Eco*RI and *Bam*HI. The ~ 2 kb fragment was then ligated to the plasmid pUC18 which had also been cut with *Eco*RI and *Bam*HI, and the construct was again transformed into TOP-10. Since the *Eco*RI and *Bam*HI sites are located within the *lacZ* gene of pUC18, blue/white screening can again be used to determine if the cassette is ligated to the plasmid. Plasmid DNA from white colonies on LB + Amp + X-gal, was digested with the enzyme *Sma*I. Since there are no *Sma*I sites in pUC18, this enzyme cuts only once in the unique site introduced by the 'tail' regions of the cross-over PCR. A blunt-ended piece of DNA containing the *cat* gene with its native promoter region

(from N. McLennan, the *cat* gene, along with its promoter was excised from pUCAT18 with *Hin*dII and *Ecl*136II) was ligated to this linearised plasmid. Since the *cat* fragment was blunt-ended, and the restriction of the vector with *Sma*I resulted in a blunt-ended product, the *cat* fragment could be ligated to the vector in either orientation; colonies were selected in which the *cat* gene is transcribed in the same direction as *dapA* would have been (Fig. 3.6.7).

3.6.4 construction of MGM100ΔdapA and MGM100DapA·His / HSV

The previous section dealt with the construction of the *dapA*::*cat* cassette. The following method was used to transfer this construct onto the chromosome of *E. coli*.

In the first instance the plasmid of interest (pUCDapA::*cat*) was transformed into competent MG1655 and allowed to grow overnight at 37°C. During this time there was a crossover between the homologous regions *nlpB* and *gcvR* on the plasmid and chromosome. A P1 lysate was then made from the MG1655 and used to transduce NK6056, selecting for Cmp^R and Tet^S colonies that could grow in the absence of adenine but required DAP (NK6056 contains *purC*::*Tn*10 and in *E. coli* *purC* is within 2 minutes of *dapA*, so any colonies which were both $purC^+$ and Cmp^R after the P1 transduction and also required the presence of DAP, had the *dapA*::*cat* construct on the chromosome). By screening for Amp^S it was also possible to determine if the plasmid had resolved from the chromosome (Fig. 3.6.1, part B).

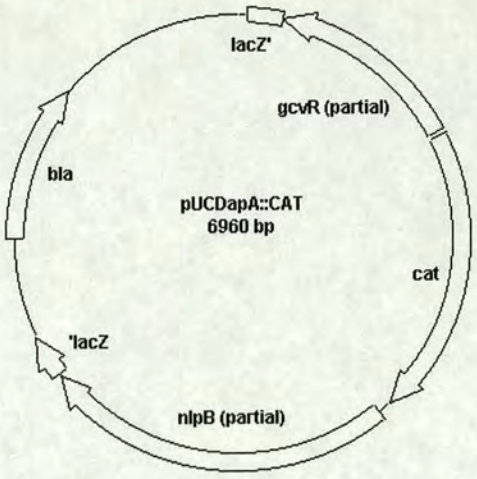


Fig. 3.6.7 The plasmid pUCDapA::CAT, where *cat* has replaced *dapA* using the technique of cross-over PCR.

Any of the transduced NK6056 colonies that were Cmp^R, Tet^S, Amp^S, $purC^+$ and required the presence of DAP were used as the donor for a new P1 lysate, which was then used to transduce MGM1655 or MGM100, selecting now only for Cmp^R and screening for DAP dependence (also for MGM100, the presence of arabinose for growth).

Surprisingly the Δ*dapA* strain would not grow on LB + DAP, although it would grow on minimal media + DAP. Though the reason for this is still unclear, it was hypothesised that since MGM100/MG1655 have a much faster growth rate in LB (generation time ~ 22 minutes) than they have in minimal media (generation time ~ 40 minutes) it was possible that the cells could not take up DAP at a sufficient rate to make new peptidoglycan when growing on LB. If this were the case then the cells would grow faster than they could make new cell wall, and would soon lyse; when on minimal media, the cells made new peptidoglycan less rapidly and could consequently match the supply of DAP with its demand.

The strain MGM100Δ*dapA* showed no enzymatic activity for DapA using the o-aminobenzaldehyde assay as described in materials and methods.

The plasmids pUCdapHIS and pUCDapHSV were able to complement MGM100Δ*dapA*, allowing the strain to grow in the absence of DAP. The presence of

either of the plasmids also allowed the strain to grow on LB + DAP as well as minimal media + DAP.

The construction of MGM100DapA·His and MGM100DapA·HSV was primarily the same as for the construction of MGM100Δ*dapA* except that the presence of DAP was not required when screening colonies.

3.7 Native structure of DapA in GroE depleted cells

Earlier in this work it had been shown that in GroE depleted cells, *dapA* is still transcribed and translated, but as the cells become more GroE depleted, the activity of the enzyme decreases. Since previous studies had only measured the amount of denatured protein in the cell, it was impossible to determine if the native structure of DapA was affected in GroE depleted cells. Since DapA is only active when it forms a homotetramer it was possible that DapA was affected by GroE depletion in such a way that it could not form a tetrameric structure.

A HSV tagged form of DapA was constructed in MGM100, which exhibited normal levels of DapA activity (i.e. an equivalent amount of protein from MGM100 and MGM100DapA·HSV gave the same level of DapA enzymatic activity). When protein extracted from this strain was run on a denaturing gel and probed using anti-HSV IgG, a monomer of the correct molecular weight was observed (Fig. 3.7.1).

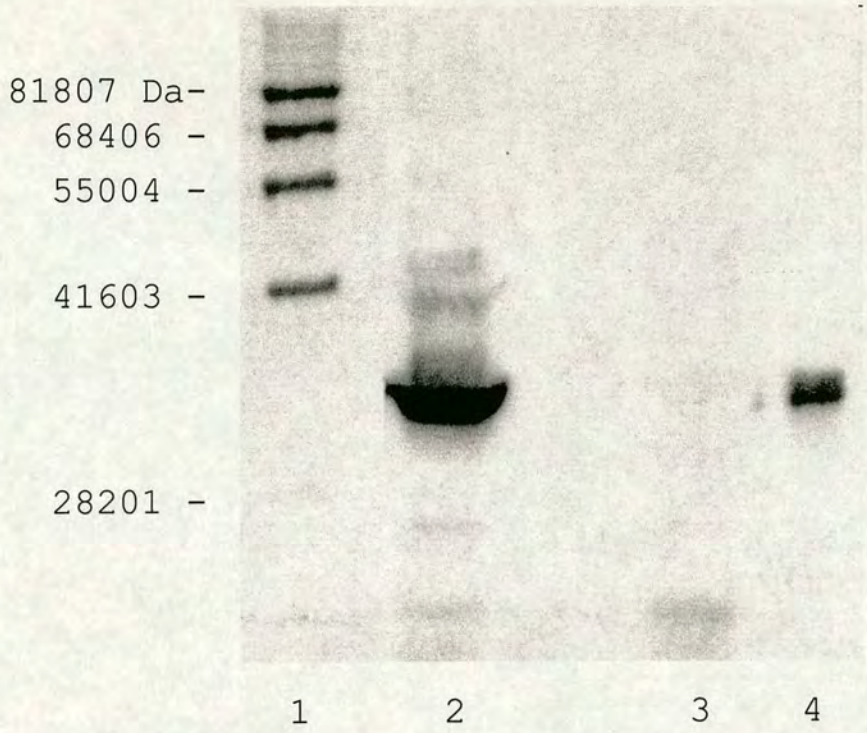


Fig. 3.7.1 Chemiluminescent detection of DapA·HSV probed using anti-HSV IgG. Lane 1: Protein molecular weight markers, lane 2: whole cell extract from MGM100DapA·HSV (DapA monomer = 31273 Da), lane 3: whole cell extract from MGM100, lane 4: positive control extract (HSV tagged protein of 31102 Da).

To determine the native structure of DapA in GroE depleted cells, MGM100DapA·HSV was grown in either LB + ara ($GroE^+$) or glu ($GroE^-$) and samples were taken at various times and analysed on an 8% (w:v) native gel with Western blotting and detection with anti-HSV IgG (as per the protocol in materials and methods). The results, shown in figure 3.7.2 show that as MGM100 becomes GroE depleted, DapA is still present in a tetrameric form, even though the samples show hardly any DapA activity. Also, the amount of the tetrameric form, rather than decreasing, actually increases (see lane 7 on figure 3.7.2). This is consistent with the fact that the level of transcription of DapA is regulated by the intracellular level of DAP, and therefore, as cells become GroE depleted, the amount of DapA increases (discussed in chapter 5). Results of the densitometry analysis of this Western blot can be seen in Table 3.7.1.

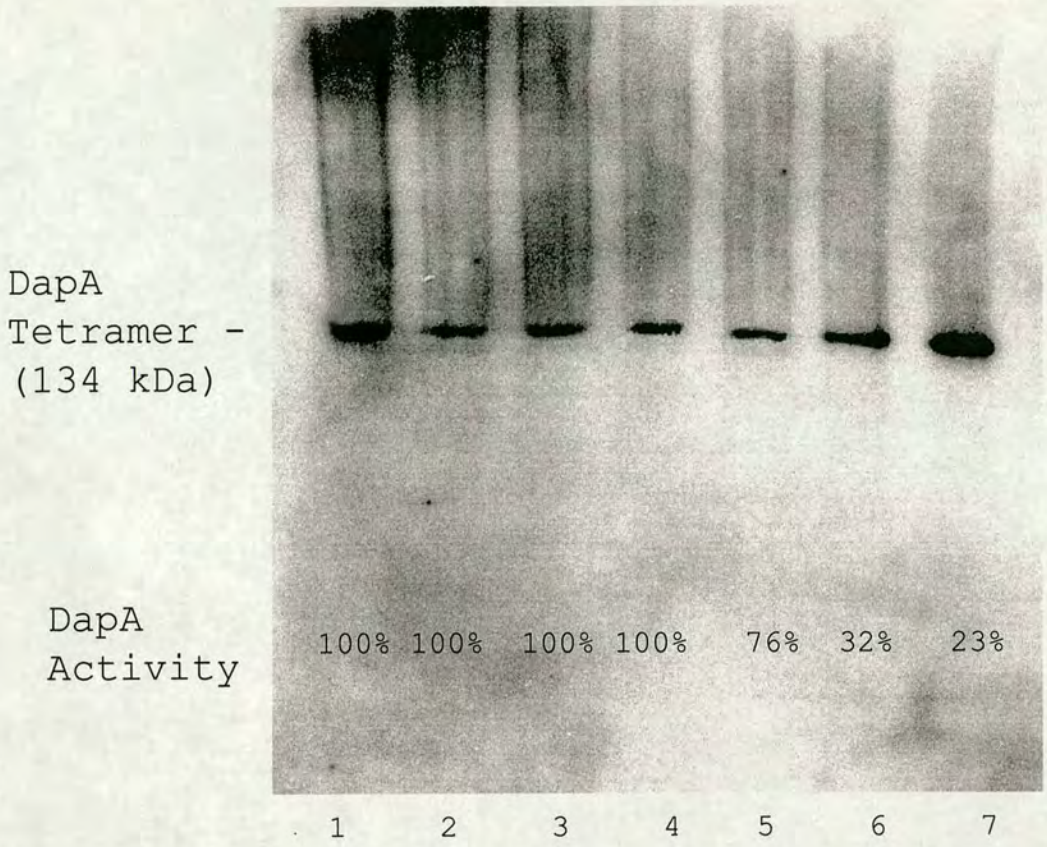


Fig. 3.7.2 Chemiluminescent detection of DapA tetramer from MGM100DapA·HSV. Lane 1: $GroE^+$ sample at T = 0 minutes after dilution into LB + ara/glu, lane 2 $GroE^+$ sample, T = 60 minutes, lane 3: $GroE^+$ T = 135 minutes, lane 4: $GroE^+$ sample, T = 150 minutes, lane 5: $GroE^-$ sample, T = 60 minutes, lane 6: $GroE^-$ sample, T = 135 minutes, lane 7: $GroE^-$ sample, T = 150 minutes. Also, lane 1 had 2x more protein loaded than any other lane.

That DapA is present in GroE depleted cells as a homotetramer, yet has no enzymatic activity could be caused by two possibilities. (i) DapA is a GroE substrate, and requires GroE to fold to a enzymatically active native state, yet in the absence of GroE DapA may still be able to fold to a near-native state such that it can form a homotetramer but has no enzymatic activity, (ii) In GroE depleted cells, DapA activity is inhibited through some unknown mechanism.

	$GroE^+$ 0 mins.	$GroE^+$ 60 mins.	$GroE^+$ 135 mins.	$GroE^+$ 150 mins.	$GroE^-$ 0 mins.	$GroE^-$ 60 mins.	$GroE^-$ 135 mins.	$GroE^-$ 150 mins.
Densitometry reading	64992 (32469)	36128	42370	41206	64992 (32469)	34379	65932	108021

Table 3.7.1 Densitometry readings of the amount of native DapA in wild type and GroE depleting cells. The Western blot in figure 3.7.2 was analysed by densitometry, the analysis shows that in a $GroE^+$ cell the level of native DapA remains approximately the same at all time points, while in a GroE depleting cell, the amount of native DapA increases as the cell becomes GroE depleted. (The 0 minute sample had two times more protein loaded than any other sample, the first densitometry reading is the actual value from the Western, while the figure in brackets is half of that value, i.e. the densitometry reading if only 1x protein had been loaded).

3.8 DapA is not inhibited in GroE depleted cells

Since DapA has no enzymatic activity in GroE depleted cells it is possible that during GroE depletion, an inhibitor molecule for DapA is synthesised in MGM100. There are many modes of inhibition in *E. coli*, for example the inhibitor molecule could have a similar structure to aspartate semialdehyde (the substrate for DapA) to the extent that it specifically binds to the active binding site, yet differs from the true substrate in that it is unreactive.

To determine if an inhibitor molecule for DapA was synthesised during GroE depletion a culture of MGM100 was grown in LB + ara/glu for 3 hours (in the glucose sample, the DapA activity was >25% of the arabinose sample) and then 20 ml of OD_{600} 0.1 cells were removed. These cells were centrifuged at 1500 x g for 5 minutes and the supernatant discarded. The cells were then resuspended in 500 µl of 0.1 M Tris buffer (pH 7.4) and sonicated to lyse the cells. If there were an inhibitor present in the GroE depleted cells, then, in theory, it should be present in the $GroE^-$ lysate. 100 nm of purified DapA was then mixed with 200 µl of either the $GroE^+$ or $GroE^-$ lysate, and allowed to incubate at 25°C for 15 minutes, then the DapA activity of the $GroE^{+/-}$ lysates (200 µl), the $GroE^{+/-}$ lysates with 100 nm DapA (200 µl) and 100 nm of DapA were measured.

For the $GroE^+$ lysate incubated with DapA, the final DapA activity should be the sum of the activity of the lysate and the purified DapA. For the $GroE^-$ lysate

incubated with DapA, if DapA activity is inhibited, the final activity should be less than the sum of the activity of the lysate and the purified DapA. The results are shown in Table 3.8.1.

The data suggests that incubating DapA with a $GroE^-$ lysate does not result in any greater decreased DapA activity than incubating DapA with a $GroE^+$ lysate. From this it can only be concluded that, under these circumstances, $GroE^-$ cells do not seem to contain a diffusible molecule that is inhibitory to DapA activity. It is possible that any inhibitor molecule is covalently attached to DapA in the $GroE^-$ lysate, and therefore was not present as a free compound in the soluble fraction.

	(i) DapA activity of lysate	(ii) DapA activity of lysate + 100nm DapA	(iii) DapA activity of 100 nm pure DapA	Percentage activity of 100 nm DapA after incubation with lysate (ii-i/iii)*100
$GroE^+$	2	2.27	0.35	77%
$GroE^-$	0.73	1	0.35	77%

Table 3.8.1 Results of inhibition experiment of DapA. DapA activity is measured as arbitrary units. After incubation with lysates from both $GroE^+$ and $GroE^-$ cells, DapA activity is the same, suggesting that a $GroE^-$ lysate contains no more inhibitory material than a $GroE^+$ lysate.

3.9 Purification of DapA from GroE depleted cells

Since DapA did not appear to be inhibited in GroE depleted cells, it was possible that the decrease in DapA activity was due to a structural problem. The native gel analysis shows that DapA can form a homotetramer in conditions where GroE is absent, but this method will only show gross structural changes and is not sensitive enough to show a slight structural change, for example in the catalytic site, that might cause the enzyme to lack activity. It was hypothesised therefore that DapA is a GroE substrate and requires the presence of GroE to fold correctly to its active form, yet in the absence of GroE may be able to fold to a near-native state, but will not fold one hundred percent correctly, and will consequently have no enzymatic activity.

To determine if this were indeed the case it was necessary to purify DapA from $GroE^+$ and $GroE^-$ cells and determine the structure of the protein using a number of methods, including circular dichroism (van Stokkum et al, 1990), nuclear magnetic resonance spectroscopy (Ferentz et al, 2000), Fourier transform infrared spectroscopy (Surewicz et al, 1993) and microcalorimetry (Cooper, 1999).

In an attempt to purify DapA from $GroE^{+/-}$ cells, 1 litre of the strain MGM100DapA·HIS was grown in LB + ara or glu for 135 minutes, the entire culture was centrifuged at 1500 x g for 10 minutes in a chilled rotor and the supernatant was

discarded. The cell pellet was then resuspended in 5 ml of His-Bind buffer and the protein was extracted using a His-Bind Quick 900 cartridge using the protocol given in materials and methods. While the protein that was extracted from the $GroE^-$ cells had little activity (>25% of the $GroE^+$ sample) there was very little protein in either sample, the $GroE^+$ sample producing 0.37 mg of DapA and the $GroE^-$ sample producing 0.2 mg. As 5-8 mg of DapA protein would be required to analyse the structure of the protein (requiring 25-40 litres of culture) it was decided that purification from a strain overproducing DapA would be more feasible.

The next approach was to use the plasmid pUCDap·His in MGM100Δ*dapA* to overproduce DapA under the control of its native promoter. The strain MGM100Δ*dapA* was used as the presence of the plasmid complements the *dapA*::*cat* mutation, and therefore cell viability is an easy way of selecting for the presence of the plasmid. Unfortunately, when MGM100Δ*dapA* pUCDap·His was grown in LB + glucose, the cells did not lyse (Fig. 3.9.1). As discussed in the introduction to this thesis, McLennan *et al.* found that the presence of *dapA* on a multicopy (pBR322) plasmid delayes lysis in GroE depleted cells. It was hypothesised that a 'pool' of DapA was present, folded correctly before the cell became GroE depleted, and this could delay the onset of lysis. Following this logic, a pUC based plasmid, having a higher copy number still would have an even larger 'pool' of correctly folded DapA. When MGM100 becomes GroE depleted, the cells eventually stop growing, if this happens before the pool of correctly folded DapA has been depleted, the cells will never lyse. This, therefore, renders the plasmid pUCDap·His useless in the purification of DapA from GroE depleted cells, as it will not be possible to extract a homogenous form of DapA, there will always be a mixture of correctly and incorrectly folded DapA.

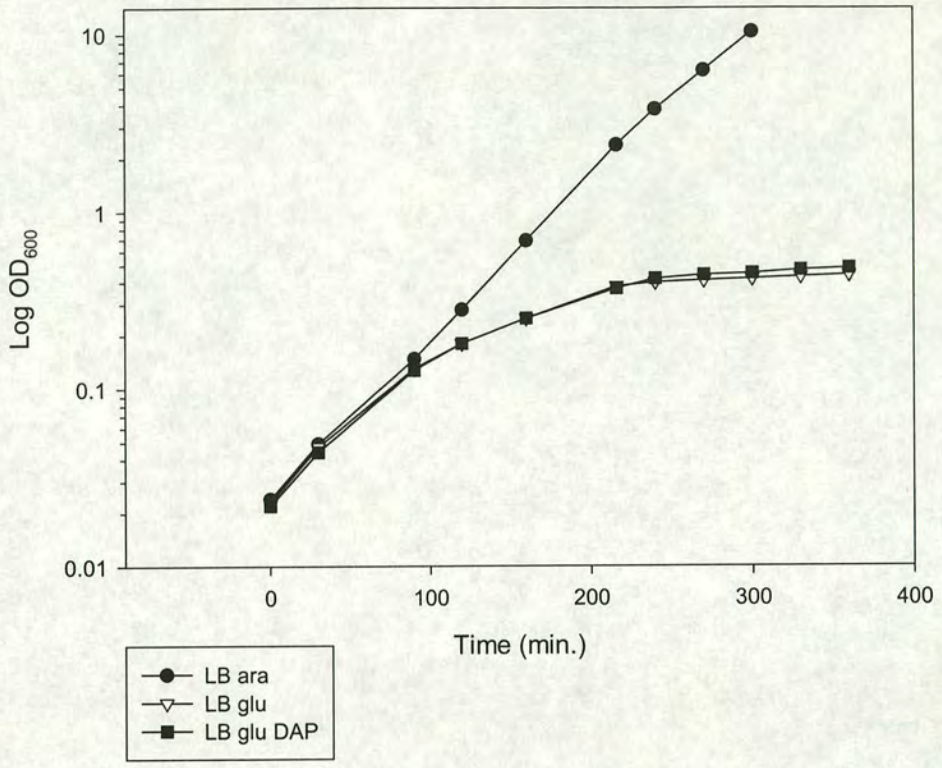


Fig. 3.9.1 Growth curves of MGM100Δ*dapA* pUCDap·His in LB + ara ($GroE^+$), glu ($GroE^-$) or glu DAP ($GroE^-$, no lysis). For clarity, when cells have been diluted 1/10, the OD has been multiplied by 10 and plotted on the graph. A 1/10 dilution was performed on all samples at 90 minutes and on the ara sample at 210 minutes.

3.9.1 Construction of pT7-5DapA·His

One of the over-expression systems that was used to try and purify DapA from GroE depleted cells was the T7 system. The pT7 range of vectors contain the T7 RNA polymerase promoter ϕ10 in front of a polylinker sequence (Tabor et al, 1985). When the gene of interest (*dapA·His*) is cloned into the multiple cloning site of the vector, its expression is regulated by the amount of T7 RNA polymerase in the cell. The strain MGM100λDE3 [constructed using the λDE3 lysogenisation kit (Novagen)] has a chromosomal copy of the T7 RNA polymerase gene under the control of a *lacUV5* promoter, thereby making the expression of DapA·His dependent on the addition of the gratuitous inducer IPTG.

Two primers were designed (UP-DAPA and DWN-DAPA) which, when used in a PCR with MGM100DapA·His chromosomal DNA as the template, would generate a product of ~1 kb which would contain the coding sequence for the His tagged version of *dapA*, with its native ribosome binding site, but without its –10 and –35 region. The primers would also introduce engineered sites into the PCR product such that it would have an *Eco*RI site at its 5′ end and a *Hin*dIII site at its 3′ end.

UP-DAPA: 5′ TGCTTGCTTTGAATTCCATACCAAAC 3′
*Eco*RI

DWN-DAPA: 5′ GAGCTGAATGAAGCTTTACCAAACGA 3′
*Hin*dIII

The PCR product was sub cloned into pGEM-T, and then released from the vector by cutting with the restriction enzymes *Eco*RI and *Hin*dIII. The restriction was run on an agarose gel, the 1 kb band corresponding to *dapA·His* was excised and cleaned up using the β-agarase protocol as described in materials and methods. The fragment was then ligated to the plasmid pT7-5, which had also been cut with *Eco*RI and *Hin*dIII. The final product, pT7-5DapA.His (Fig. 3.9.2) was then transformed into the strain MGM100λDE3.

The strain was grown in LB + ara/glu and IPTG was added (0.4 mM) at 150 minutes. After 30 minutes the samples were run on an SDS-PAGE gel according to the protocol given in materials and methods. In the $GroE^+$ cells there was a clear increase of *dapA·His*, but in $GroE^-$ cells there was no increase in *dapA·His* (Fig. 3.9.3).

This experiment was repeated a number of times, with both varying concentrations of inducer and adding the inducer at varying times (though no earlier than 120 minutes, as the $GroE^-$ cells would not be sufficiently GroE depleted to produce misfolded DapA); in $GroE^-$ cells there was consistently no induction of *dapA·His*.

Since the list of verified GroE substrates is far from complete, it was hypothesised that T7 RNA polymerase was affected by GroE depletion causing the lack of expression using this system in GroE depleted cells.

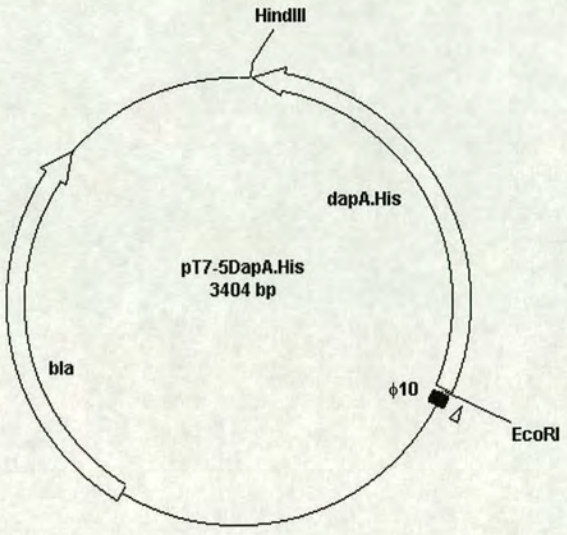


Fig. 3.9.2 The plasmid pT7-5DapA·His. The His tagged version of the gene *dapA* is cloned into the multiple cloning site of the plasmid, putting it under the control of the ϕ10 promoter.

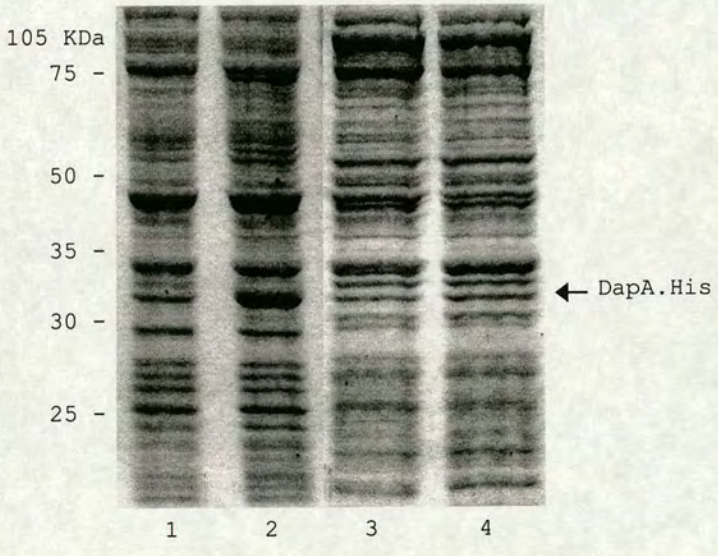


Fig. 3.9.3 Coomassie blue stained SDS-PAGE gel of protein from MGM100λDE3 pT7-5DapA·His. Lane 1: $GroE^+$ before induction, lane 2: $GroE^+$ 30 min. after induction (0.4mM IPTG), lane 3: $GroE^-$ before induction, lane 4: $GroE^-$ 30 min. after induction (0.4mM IPTG). The induced band corresponding to DapA·His (31.2 KDa) is clearly visible in lane 2.

3.9.2 Construction of pJFDapA·His

To overcome the problems of the T7 system the His tagged *dapA* gene was cut out of pT7-5DapA·His using the restriction enzymes *Eco*RI and *Hin*dIII and ligated to the plasmid pJF118(EH) (Furste et al, 1986) which had also been cut with *Eco*RI and *Hin*dIII . The resulting plasmid, pJFDapA·His had the *dapA·His* gene under the direct control of a *tac* promoter (Fig. 3.9.4). The plasmid also carried the $lacI^{Q}$ gene which prevents transcription from the *tac* promoter in the absence of the inducer IPTG and therefore tightly regulates transcription of *dapA·His*.

The plasmid was transformed into competent MGM100 and the cells were grown in LB + ara/glu as per the instructions in section 3.9.1. IPTG was added to a final concentration of 0.4 mM and the samples analysed on an SDS-PAGE gel according to the protocol in materials and methods.

The results (Fig. 3.9.5) again show that while DapA·His is over-expressed in the $GroE^{+}$ sample, it is not over-expressed in the GroE depleted sample. It was at this point that it was discovered that LacI is affected by GroE depletion. In GroE depleted cells, LacI does not release from LacO when IPTG is present (discussed in detail in chapter 4), therefore nothing is expressed from a *lac* or *tac* promoter in GroE depleted cells. This would also explain why the T7 vector failed to overproduce any protein. The T7 RNA polymerase required to transcribe the mRNA of the gene under the control of the $\phi 10$ promoter, was itself under the control of a *lacUV5* promoter, and was consequently never produced in a GroE depleted cell.

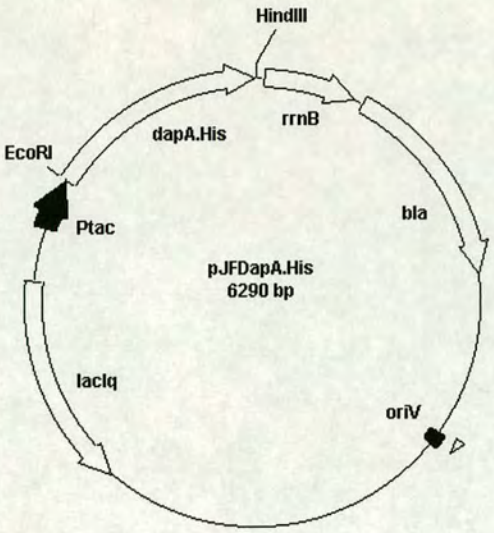


Fig. 3.9.4 The plasmid pJFDapA·His. The His tagged version of the *dapA* gene was cut out of the plasmid pT7-5DapA·His using the restriction enzymes *Eco*RI and *Hin*dIII and cloned into the *Eco*RI and *Hin*dIII sites of pJF118(EH).

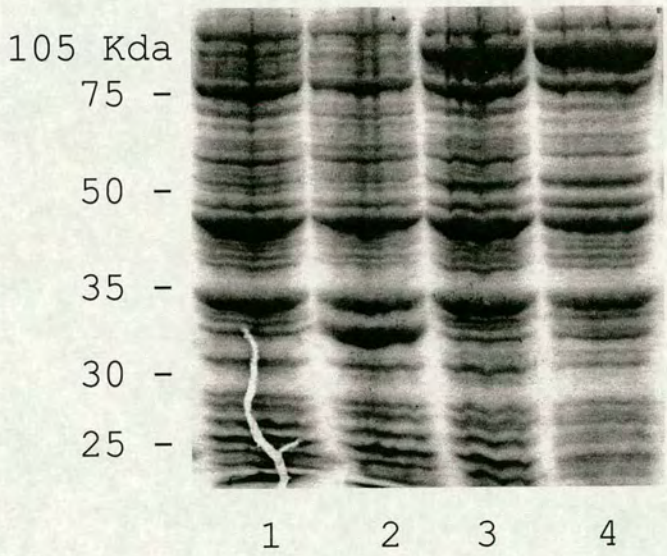


Fig. 3.9.5 Coomassie blue stained SDS-PAGE gel of protein from MGM100 pJFDapA·His. Lane 1: $GroE^+$ before induction, lane 2: $GroE^+$ 60 min. after induction (0.4mM IPTG), lane 3: $GroE^-$ before induction, lane 4: $GroE^-$ 60 min. after induction (0.4mM IPTG). The induced band corresponding to DapA·His (31.2 KDa) is clearly visible in lane 2.

3.9.3 Construction of pPROTet.E132DapA·His

Since it was not possible to use a *lac* promoter based system to overproduce protein in GroE depleted cells, the PRO bacterial expression system was used instead (Clontech). The PROTet.E vectors utilise a novel tetracycline-regulated promoter, $P_{Ltet0-1}$ that takes advantage of the high expression levels from the P_L promoter of phage λ and the tight control provided by operator 2 of the Tn*10* tetracycline resistance operon (Lutz et al, 1997). Thus, $P_{Ltet0-1}$ is tightly repressed by the highly specific Tet repressor protein and induced in response to anhydrotetracylcine (aTc) allowing control of induction over a wide range. (Anhydrotetracycline is a derivative of tetracycline that acts as a more potent inducer of PROTet.E systems and is less toxic to *E. coli* than tetracycline). $P_{Ltet0-1}$ has been paired with the Col E1 high-copy replication origin (ori) to deliver a 2500-fold range of inducibility. At full induction levels, the target protein can represent as much as 10% of total protein. The Spc^R plasmid pPRO, used in conjunction with the PROTet.E plasmids, expresses defined amounts of the Tet repressor to ensure reliable intracellular conditions.

The His tagged version of dapA was cloned into the vector pPROTet.E132 in the following manner; the *dapA·His* gene was amplified using PCR on chromosomal DNA from MGM100DapA·His, the primers used were DWN-DAPA (from section 3.9.1) which introduced a 3′ *Hin*dIII restriction enzyme recognition site, and the primer UPDAPA-K, which introduced a *Kpn*I recognition site into the 5′ end.

UPDAPA-K: 5′ TGCACAGAGGGGTACCCATGTTCACG 3′
*Kpn*I

The important aspect of this primer was that it introduced the *Kpn*I site one base before the ATG of *dapA·His*. Therefore when it was cloned into pPROTet.E132 the number of bases between the ribosome binding site and the ATG of the inserted gene would be conserved.

pPROTet.E132:

5′ .....**ATTAAAGAGGAGAAA**GGTACC**ATG**GGTGAA..... 3′
$P_{Ltet0-1}$ RBS *Kpn*I Myc-Tag *Hin*dIII

5′ end of *dapA·His* after PCR with DWN-DAPA and UPDAPA-K:

```
5′ .........TGCACAGAGGGGTACCCATGTTCACG.......... 3′
                          KpnI   dapA·His
```

Therefore when *dapA·His* was cloned into the *Kpn*I and *Hin*dIII sites of pPROTet.E132, the initiation codon of *dapA·His* would be in the same position as the initiation codon of the Myc Tag would have been, thereby assuring that both the $P_{Ltet0-1}$ promoter and ribosome binding site were used to give the best level of expression.

The *dapA·His* fragment was amplified by PCR, run on an agarose gel and the 1 kb fragment extracted using the β-agarase protocol in materials and methods. The fragment was then sub cloned into pGEM-T, and then released by cutting with the restriction enzymes *Kpn*I and *Hin*dIII. The fragment was then ligated to pPROTet.E132 which had also been cut with *Kpn*I and *Hin*dIII. The resulting plasmid pPROTet.E132DapA·His (Fig. 3.9.6) now had *dapA·His* expression under the control of the promoter $P_{Ltet0-1}$.

The plasmid was then transformed into MGM100PRO (MGM100 containing the Spc^R pPRO plasmid which carries the *tetR* gene) and transformants were selected on LB + Kan, Spc, Cmp and 0.2% ara.

To test if DapA·His was produced in GroE depleted cells using the $P_{Ltet0-1}$, the cells were grown in LB + ara/glu using the method described in section 3.9.1 and the inducer, aTc, was added to a final concentration of 100 ng ml^{-1}, with samples being taken before and after induction for analysis on an SDS-PAGE gel. The results (Fig. 3.9.7) show that using the PROTet.E system, DapA·His can be expressed in GroE depleted cells.

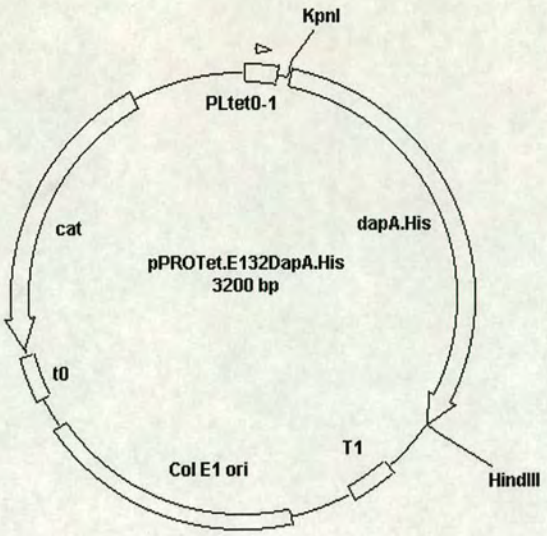


Fig. 3.9.6 The plasmid pPROTet.E132DapA·His. The *dapA·His* gene, being cloned into the *Kpn*I and *Hin*dIII sites of pPROTet.E132 is now under the control of the promoter $P_{Ltet0-1}$.

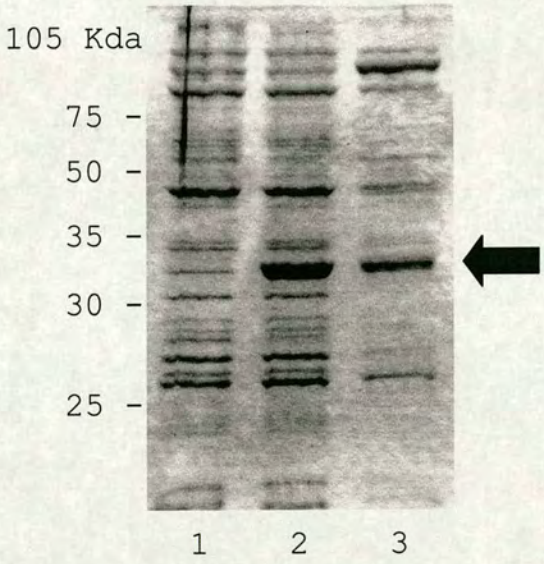


Fig. 3.9.7 Coomassie blue stained SDS-PAGE gel of protein from MGM100PRO pROTet.E132DapA·His. Lane 1: $GroE^+$ before induction, lane 2: $GroE^+$ 60 min. after induction (100 ng ml^{-1} aTc), lane 3: $GroE^-$ 60 min. after induction (100 ng ml^{-1} aTc). The induced band corresponding to DapA·His (31.2 KDa) is clearly visible in both lanes 2 and 3.

3.10 Over-expressed DapA is insoluble in GroE depleted cells

The plasmid pROTet.E132DapA·His was used to express DapA·His in GroE depleted MGM100PRO, the cells were then lysed and the protein extracted from the soluble fraction of the lysate using the protocol described in materials and methods. When the elution step was performed in the purification protocol it was found to contain no DapA·His, yet performing the same procedure on $GroE^{+}$ cells resulted in a high yield of DapA·His. Since DapA·His was being expressed in these cells (as determined by SDS-PAGE analysis) it is unclear why the protein cannot be purified.

It was then observed that when expression of DapA·His was induced in MGM100PRO the cells formed large inclusion bodies (Fig. 3.10.1) suggesting the possibility that the over-expressed protein was aggregating. When a protein aggregates it is usually found in the insoluble fraction of an *E. coli* cell lysate, this might explain why the protein could not be extracted from the soluble fraction of a GroE depleted cell lysate.

To determine if DapA was becoming insoluble when over-expressed in GroE depleted cells, MGM100PRO pROTet.E132DapA·His was grown in LB+ glu/ara and induced with aTc (100 ng ml^{-1}) at 150 minutes after the dilution into LB glu or LB ara. Samples were taken after induction and one of two procedures was performed: either the cells were resuspended in 1 x PAGE loading buffer and lysed by boiling for 3 minutes (this would result in a sample which contained both soluble and insoluble proteins) or the samples were resuspended in 0.1 M Tris buffer and sonicated to lyse the cells. The samples were then centrifuged (10 minutes at 14000 x g) to pellet insoluble material and the soluble fraction mixed with 1 x PAGE loading buffer (resulting in a sample containing only soluble proteins). The samples were run on an SDS-PAGE gel and analysed.

The results (Fig. 3.10.2) required careful analysis as the soluble fraction sample always contained less protein than the soluble + insoluble fraction sample. By choosing a control band (the upper band in the gel image) and normalising all lanes so that they had the same intensity for that band, the relative intensities of the band representing DapA·His could be determined.

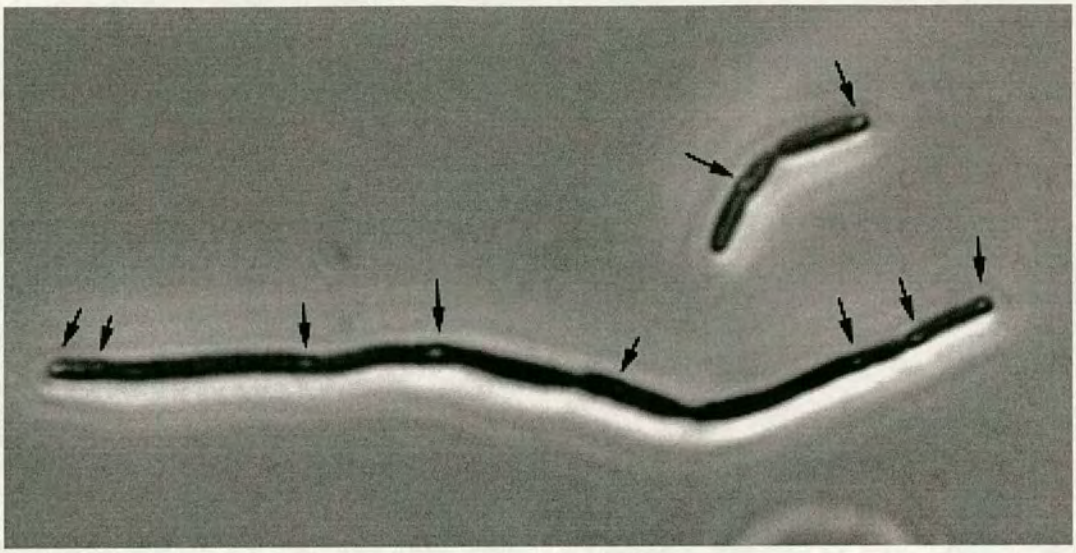


Fig 3.10.1 Inclusion bodies in MGM100PRO over-expressing DapA·His under $GroE^-$ conditions.

To determine if the control band changed intensity in GroE depleted cells the intensity of the DapA·His band from the $GroE^+$ sample (soluble + insoluble fraction) and $GroE^-$ sample (soluble + insoluble fraction) were compared when the intensities had been normalised using the control band. If the control band was affected by GroE depletion (either increasing or decreasing in intensity) then the values for the DapA·His band from the two samples should be quite different. The value of the DapA·His band from the $GroE^+$ sample was 1760 (arbitrary units) and the value from the $GroE^-$ sample was 1648. This suggested that the control band was unaffected by GroE depletion and was therefore acceptable to normalise against.

The $GroE^+$ sample (soluble fraction) had a DapA·His value of 1101.5 when normalised, while the $GroE^-$ sample (soluble fraction) had no band corresponding to DapA·His. This data suggests that even in a $GroE^+$ sample, some fraction (in this instance, ~ 37%) of over-expressed DapA·His is insoluble, but in GroE depleted cells 100% of DapA·His is the insoluble.

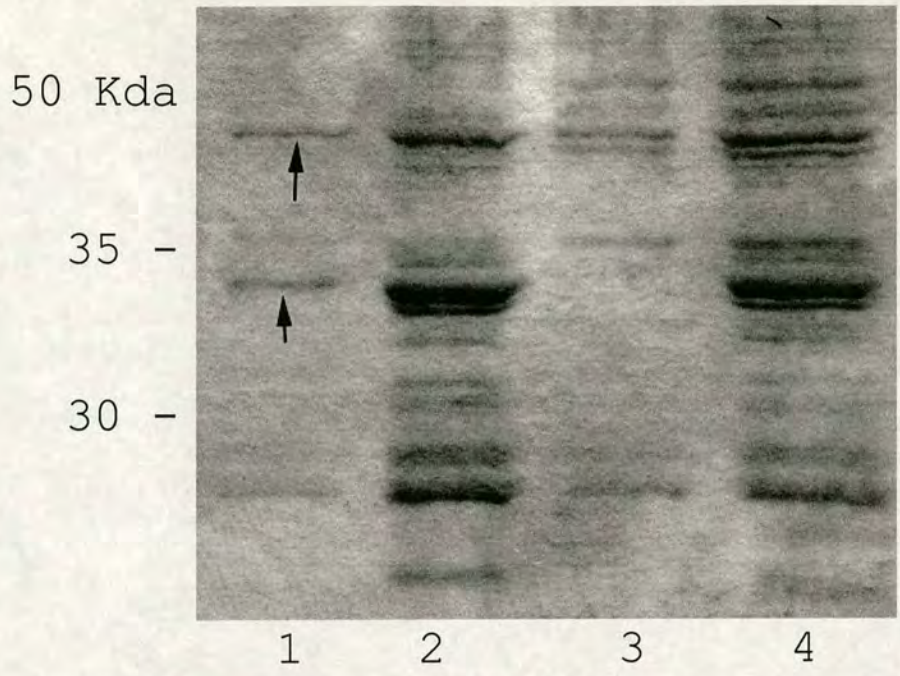


Fig. 3.10.2 Coomassie blue stained SDS-PAGE gel of proteins from MGM100PRO pROTet.E132DapA·His induced with aTc (100ng ml^{-1}). Lane 1: $GroE^+$ soluble fraction, lane 2: $GroE^+$ soluble + insoluble fraction, lane 3: $GroE^-$ soluble fraction, lane 4: $GroE^-$ soluble + insoluble fraction. The upper arrow in the image is pointing to the control (normalising) band and the lower arrow is pointing to the band representing DapA·His.

It is not surprising that even in the $GroE^+$ sample some of the over-expressed DapA·His is aggregating. If, as it is hypothesised, DapA is a GroE substrate, then when DapA is produced at a much higher level than is physiologically usual, the cell may not have enough GroE to fold all of it correctly. The incorrectly folded protein would then act like DapA from a $GroE^-$ cell, and consequently form aggregate.

To try and overcome the problems of aggregation of over-expressed DapA·His in GroE depleted cells, a range of aTc concentrations was used, with values from 0.1 ng ml^{-1} to 100 ng ml^{-1}. The cells were also allowed to express at different temperatures; 4°C, 22°C, 30°C and 37°C (if expression is slowed down, i.e. by growing the cells at a lower temperature, then the proteins may not aggregate) but the DapA·His protein was always found in the insoluble fraction.

The plasmid pPROTet.E132DapA·His has a copy number of 50-70 molecules per cell, yet the plasmid pBRA6 (*dapA* on pBR322) while having the same copy number does not produce inclusion bodies in MGM100. The major difference between these two plasmids is that pBRA6 contains the *dapA* gene under the control of its own promoter.

The His tagged version of *dapA* was amplified using the primers DWN-DAPA (introducing a 3′ *Hin*dIII site) and DAPANP (introducing a 5′ *Eco*RI site), the PCR product was 1.1 kb and contained the coding region for *dapA·His* with the native *dapA* promoter.

DAPANP: 5′ TGGTGTTCAGAATTCCAGGGCGAT 3′
*Eco*RI

The 1.1 kb PCR product was sub cloned into pGEM-T, released by cutting with the restriction enzymes *Eco*RI and *Hin*dIII, and then cloned into pBR322 which had also been cut with *Eco*RI and *Hin*dIII.

pBR322 is Amp^R and Tet^R but when the PCR product was cloned into the *Eco*RI and *Hin*dIII sites of the plasmid it disrupted the gene responsible for Tet^R, therefore recombinant DNA could be selected for by finding transformants that were now Amp^R and Tet^S.

The plasmid was transformed into MGM100 and the cells were grown in LB + ara/glu DAP with samples being harvested after 5 hours. Since DapA·His was being constitutively expressed in this strain, it was necessary to harvest the cells at a later time to be sure that only misfolded DapA·His was present. While pBRA6 delayed lysis in GroE depleted cells, it did not prevent it altogether, and even with pBRA6 present, GroE depleted cells lysed after 5 hours, this represented the time at which there was no longer any correctly folded DapA to prevent the cells from lysing.

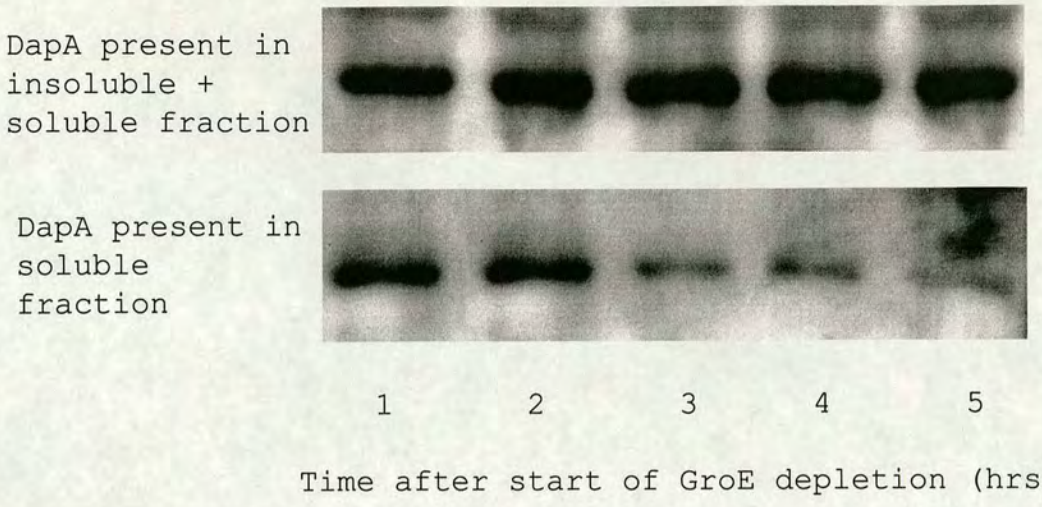


Fig. 3.10.3 Levels of DapA, measured by Western blotting and chemiluminescent detection, from the soluble and insoluble fraction of MGM100 pBRA6 at various times after the start of GroE depletion.

The DapA·His protein was then extracted according to the protocol described in materials and methods, but there was no protein detected in the final elution, though when the same procedure was performed on $GroE^+$ cells a large amount of active DapA·His was extracted. This data suggests that even though the *dapA* promoter may not be as strong as the $P_{Ltet0\text{-}1}$ promoter, and consequently not as much DapA·His is produced (the absence of visible inclusion bodies indicates that less aggregate is being made), the amount of DapA·His that is produced is still going into the insoluble fraction of *E. coli* proteins.

Figure 3.10.3 shows the results of a Western analysis of DapA levels in the soluble and insoluble fractions of MGM100 pBRA6 at various times after GroE depletion. The data clearly shows that, as the cells become more GroE depleted, less DapA protein is found in the soluble fraction.

3.11 Summary and discussion

Previous work on DapA in *E. coli* had suggested that levels of the DapA polypeptide were greatly decreased in GroE depleted cells (McLennan et al, 1998). It was hypothesised that in the absence of the chaperone GroELS, DapA was misfolded and therefore became a target for proteolysis; while DapA was being actively translated in GroE depleted cells the protein had a rapid turnover rate.

In this work pulse chase experiments were performed on DapA: newly synthesised protein were labelled by incorporating radioactive methionine residues into the protein and then an excess of non-radioactive methionine was added. The addition of an excess of non-radioactive methionine effectively quenched the incorporation of any more radioactive methionine. The rate of decay of the radioactive proteins could then be measured to determine their stability. The proteins of both $GroE^-$ and $GroE^+$ *E. coli* were separated using 2-D electrophoresis and the intensity of the spot representing DapA was measured. Using this method it was shown that DapA from both $GroE^+$ and $GroE^-$ cells had a similar turnover rate, suggesting that DapA levels do not decrease in GroE depleted cells.

One of the major limitations of this work is that the substrates of GroE in *E. coli* are still largely unknown, therefore the choice of a 'control' protein to compare turnover rate against is crucial. If the 'control' protein is itself affected by GroE depletion then the data from such experiments will be incorrect. In the comparison of turnover rates of DapA a number of 'control' proteins were therefore chosen to minimise the possibility of this occurrence.

When the levels of DapA were measured directly in GroE depleted cells using immunodetection they were not found to decrease as had been previously reported. In the past DapA levels had been measured relative to a control protein FtsZ, but in this work it was shown that FtsZ levels increase in GroE depleted cells. FtsZ is under the control of a σ^S promoter (Vicente et al, 1998). When cells become GroE depleted growth slows and the cells eventually stop growing. This causes an increased level of transcription from the σ^S promoter thus producing higher levels of FtsZ. Therefore by normalising FtsZ levels, rather than observing a stable level of DapA and an increased level of FtsZ the reverse is observed: a stable level of FtsZ and a decreased level of DapA.

While it had now been shown that DapA levels were constant in GroE depleted cells, the activity of the protein did decrease markedly with GroE depletion. Since the native structure of DapA is a homotetramer (Shedlarsky et al, 1970; Richaud et al, 1986) and previous measurements of DapA levels had been of the

denatured protein (i.e. it was only measuring the amount of monomeric protein) it was possible that DapA was not forming a tetramer in GroE depleted cells, which would explain the lack of activity. By extracting DapA under non-denaturing conditions and running on a non-denaturing polyacrylamide gel, the native structure of DapA could be observed. It was found that DapA from GroE depleted cells ran at the correct size for a homotetramer, much like DapA from a $GroE^+$ cell. This then suggested that in GroE depleted cells DapA can either fold well enough to form a homotetramer, but not completely correctly (having no enzymatic activity) or in GroE depleted cells DapA activity is somehow inhibited.

The ability of chaperone requiring proteins to fold to a near native state is not unheard of. The *E. coli* chaperone SecB will not bind to either the native or unfolded state of the protein barstar, but rather, binds to a late near-native intermediate along the folding pathway (Panse et al, 1998).

To determine if DapA were being inhibited in GroE depleted cells, pure DapA was mixed with a lysate containing the soluble fraction of GroE depleted *E. coli*. If there were an inhibitor molecule in the $GroE^-$ lysate then it may cause decreased activity of added purified DapA. While a small decrease in DapA activity was noted after it was incubated with a $GroE^-$ lysate, it was also noticed when the protein was incubated with a $GroE^+$ lysate. This may have been due to the presence of proteolytic enzymes in the lysates, but does not indicate the presence of an inhibitor of DapA specific to the $GroE^-$ lysate. It must also be noted that this experiment does not preclude the possibility of an inhibitor for DapA being present in a GroE depleted cell, only that inhibition of DapA was not observed under these circumstances (which are obviously quite different from *in vivo* conditions). There is also the possibility that the inhibitor binds irreversibly to DapA, and therefore in the $GroE^-$ lysate, all the inhibitor may already be bound to DapA, therefore the inhibitor would not be available to inhibit the DapA which was added during the assay.

As it did not seem that DapA had lost activity due to inhibition, an attempt was made to purify DapA to determine if there were indeed minor structural differences between DapA from $GroE^+$ and $GroE^-$ cells. To achieve this a His or HSV tagged form of DapA was constructed and used to replace the wild-type DapA of MGM100. The tags were not found to cause any change in the activity of the protein (and therefore did not affect the correct folding of the protein) yet they provided a way to either detect the protein using anti-His or anti-HSV antibodies or to purify the protein by affinity chromatography. Unfortunately the purification strategy of DapA from GroE depleted cells met with a number of obstacles. Firstly, it was found that *lac* based over-expression vectors did not work in GroE depleted cells

(this phenomena is described in greater detail in chapter 4). This again further illustrates one of the major problems with the search for GroE substrates; as the substrates for GroE in *E. coli* are largely unknown, it is almost impossible to predict what will be affected by GroE depletion, as in this case where a promoter used to over-express a putative GroE substrate was itself affected by GroE depletion!

The second obstacle in the purification of DapA in GroE depleted cells was that the slightly misfolded DapA was highly prone to aggregation when the gene was present in more than one (chromosomal) copy. While the reasons for this behaviour remain unknown there are a number of possibilities that could explain it. When DapA was under the control of the $P_{Ltet0\text{-}1}$ promoter (Lutz et al, 1997) it was expressed at a very high rate, much higher than when it was under the control of its native promoter. And when DapA was present on a multicopy plasmid (pBR or pUC) there were many more copies of the *dapA* gene being transcribed than if only the chromosomal copy were present. It may also be hypothesised that in the absence of GroE, DapA, while being able to fold almost correctly, would have a different folding kinetic.

Hen lysozyme does not require GroEL for folding *in vitro*, yet in the presence of GroEL an early intermediate is bound and the slow kinetic phase that reflects the reversal of non-native interactions is accelerated (Coyle et al, 1999). This observation suggests a mechanism for GroEL-assisted folding in which the reorganisation of non-native tertiary interactions is facilitated but domain folding is unperturbed. If this were the case with DapA then in the absence of GroE the newly synthesised protein may fold more slowly to its near-native state than DapA would fold to its native state in the presence of GroE. Also if there were many more copies of the slowly folding protein present in the cell (such as would be the case when *dapA* was on pBR or pUC) then there is an increased risk that exposed hydrophobic regions would interact leading to aggregation. It may therefore only be possible to purify DapA from GroE depleted cells using a His-tagged form on the chromosome and extracting the protein from large batches of culture.

Although DapA seems to have a requirement for both GroEL and GroES to refold *in vitro*, the data from the experiment described in section 3.3.1 suggests that only GroEL is required for folding *in vivo*. There is the possibility that enough GroES was still being made from the chromosome to allow the GroELS reaction cycle to occur, but the levels would be very low at 5 hours after the start of depletion (~0.25%) and therefore this does not seem likely.

The plasmid pGroEL has the gene *groEL* under the control of a *lac* promoter (the plasmid was used before it was discovered that expression from *lac* promoters

decreases in GroE depleted cells) on a multicopy plasmid. Theoretically when this plasmid was transformed into MGM100 and the cells grown in the presence of glucose, transcription from the chromosomal *groELS* genes would cease and transcription from the plasmid *groEL* gene would continue. Though what was probably happening was that upon addition of glucose transcription from the chromosomal *groELS* genes would cease (as expected) but as the cells became more GroE depleted, transcription from the plasmid *groEL* gene would also cease. This may explain why DapA activity did decrease (albeit at a slower rate than cells without the plasmid pGroEL) in these cells eventually. While the cells had a higher initial level of GroEL (allowing newly synthesised DapA to fold correctly and thus delaying the production of inactive tetramer) this eventually decreased so that all newly synthesised DapA was folded in the absence of both GroEL and GroES. To overcome this problem, GroEL synthesis could be put under the control of a promoter unaffected by GroE depletion (i.e. $P_{Ltet0\text{-}1}$) and the experiment repeated to determine if DapA only requires GroEL to fold *in vivo*. Again the hen lysozyme experiments of Coyle *et al.* demonstrate the folding of protein by GroEL in the absence of GroES. Other examples are reviewed by Wang and Weissman (Wang et al, 1999).

In conclusion, DapA has proved a most interesting potential GroE substrate, one that may be able to reach a near-native conformation in GroE depleted cells yet requires GroE to attain an enzymatically active state; a protein that may only require GroEL and not the GroELS complex for folding; a protein that is highly susceptible to aggregation when over-expressed in the absence of GroE, yet does not aggregate when present in only a single copy. To further understand this complex protein, attempts should be made to purify DapA from GroE depleted cells and compare it structurally to DapA from $GroE^+$ cells.

Chapter IV

Effects of GroE depletion on LacI

4.1 The *lac* operon

During the course of this work it was discovered that it is not possible to produce DapA from Lac based promoters in GroE depleted cells. This chapter describes the analysis of the Lac promoter in MGM100 during GroE depletion, and shows that LacI behaves differently in wild type and GroE depleted cells.

More than 30 years ago, Jacob and Monod (Jacob et al, 1961a) introduced the *E. coli* lactose operon as a model for gene regulation. The model persists as a cogent depiction of how a set of structural genes may be co-ordinately transcribed or repressed depending upon the concentration of metabolites in the growth medium.

The expression of the *lac* enzymes is regulated by three basic components: the promoter-operator sequence, a transcriptional inhibitor [*lac* repressor protein (LacI)], and a transcriptional activator (cyclic AMP-dependent catabolite gene activator protein: CAP) (Kolb et al, 1993). Cyclic AMP (cAMP) levels are inversely related to the availability of glucose via effects on adenyl cyclase activity (Kolb et al, 1993), at high concentrations of cAMP (i.e. low concentrations of glucose) a complex of the CAP protein with cAMP raises the amount of transcription by binding a recognition site on the DNA adjacent to RNA polymerase and thereby the affinity of RNA polymerase for the promoter is increased (Zubay et al, 1970). This arrangement ensures the most direct metabolic route to energy production through preferential utilisation of glucose under conditions where lactose is also available. The partially twofold symmetrical operator DNA target sequence for LacI overlaps the promoter sequence (Gilbert et al, 1973). RNA polymerase binding (Schlax et al, 1995), initiation (Straney et al, 1987), and/or elongation (Lee et al, 1991) are inhibited when LacI occupies this site, precluding production of the mRNA encoding the *lac* enzymes (Fig. 4.1.1). In the absence of lactose, the high affinity of LacI for the lactose operator sequence (LacO) allows production of only small quantities of *lac* mRNA

When lactose is available in the environment, the low constitutive amounts of *lac* permease transport this sugar into the cell, and the correspondingly low levels of β-galactosidase result in production of the *in vivo* inducer β-1,6-allolactose (Jobe et al, 1972). Binding of this natural inducer to the repressor elicits a conformational change that diminishes its affinity for the operator sequence without effect on binding to non-specific DNA (Riggs et al, 1970; Lin et al, 1975a). Excess non-operator DNA in the cell thereby effectively competes for binding to the protein and sequesters the repressor-inducer complex (Lin et al, 1975b; von Hippel et al, 1974; Kao-Huang et al, 1977), allowing transcription of *lac* mRNA to proceed so long as

lactose is available in the medium. When lactose levels are decreased, the intracellular store of inducer is depleted by β-galactosidase hydrolysis. Under these conditions, β-1,6-allolactose dissociates from the repressor protein, which resumes its conformation with high affinity for operator DNA, associates with LacO, and shuts down further synthesis of *lac* mRNA. The lactose regulatory cycle therefore involves association with both specific and non-specific DNA sequences, binding of sugar molecules and conformational shifts in response to these ligands.

The gratuitous inducer, isopropyl-β-D-thiogalactoside (IPTG) can be used to turn on transcription of the lactose operon (Jacob et al, 1961b) using the same set of events as β-1,6-allolactose, but IPTG is not a substrate for β-galactosidase and therefore it is not metabolised .

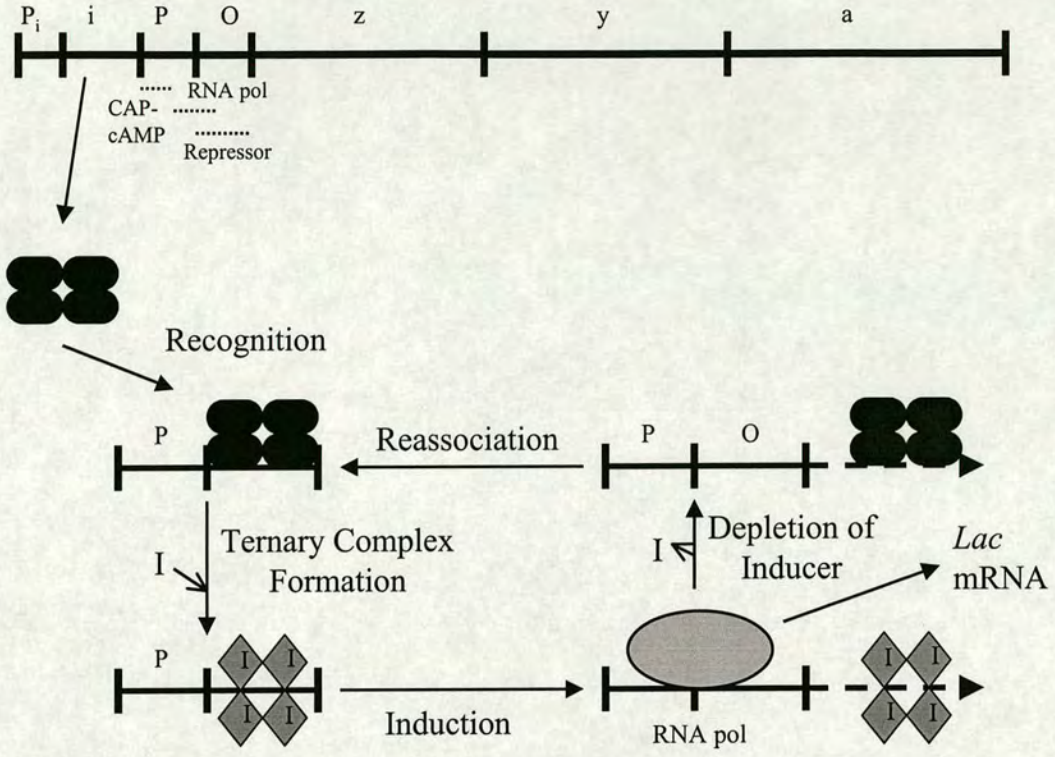


Fig. 4.1.1 Schematic of lactose operon. P_i, promoter for i gene; i, gene encoding lactose repressor (LacI); P, promoter for *lac* enzymes; O, lactose operator sequence (LacO); z, gene encoding β-galactosidase; y, gene encoding *lac* permease; a, gene encoding thiogalactoside transacetylase; I, inducer; RNA pol, RNA polymerase holoenzyme. The i gene product, lactose repressor, binds to operator to preclude transcription by RNA polymerase. In the presence of inducer sugars, LacI undergoes a conformational change that diminishes affinity for LacO but does not affect binding to non-specific DNA sites, which then compete for binding LacI-Inducer. When inducer levels are depleted, LacI resumes its state with high affinity for LacO and inhibits transcription of mRNA.

4.2 Evaluating the effect of GroE depletion on P_{Lac}

Previous observations of LacZ expression in GroE depleted cells [LacZ under the control of the *dapA* promoter, P_{DapA} (section 5.2)] has shown that LacZ folding and function is GroE independent, and therefore β-galactosidase assays can provide an accurate analysis of promoter function in GroE depleted cells. MGM100, being an MG1655 derivative, is a lac^+ strain, with LacZ under the control of the *lac* promoter on the chromosome.

Initial experiments were to grow MGM100 in buffered LB [(LB – NaCl) + (1 x VB salts)] + ara or glu + DAP (it was found that the inclusion of 1 x VB salts gave a more consistent result when performing β-galactosidase assays), and induce LacZ expression by adding IPTG at a concentration of 1mM at various stages of GroE depletion.

The results (Fig. 4.2.1) show that in GroE depleted cells LacZ expression is dramatically reduced when induced with IPTG at 1mM concentration. The two graphs show Miller Units plotted against either time or log OD_{600}. By plotting Miller Units against time it is easier to compare induction levels of samples that had different optical densities, but by plotting Miller Units against log OD it is possible to determine the amount of induction of LacZ per amount of cell growth.

Since decreased LacZ expression is not due to a failure of LacZ to fold or function or of *lacZ* mRNA to be translated, it must be due to a decrease in the level of expression of the *lac* promoter. Two possible explanations for this result are that IPTG was not entering the cell and thus was unable to promote induction, or that LacI activity had been affected either directly or indirectly by GroE depletion and was unresponsive to IPTG. An example of a direct effect on LacI by GroE depletion could be that LacI protein structure changes, which causes it to become unresponsive to IPTG, whereas an example of an indirect effect on LacI by GroE depletion might be that DNA topology could be affected and consequently LacI may remain bound to LacO, even in the presence of IPTG.

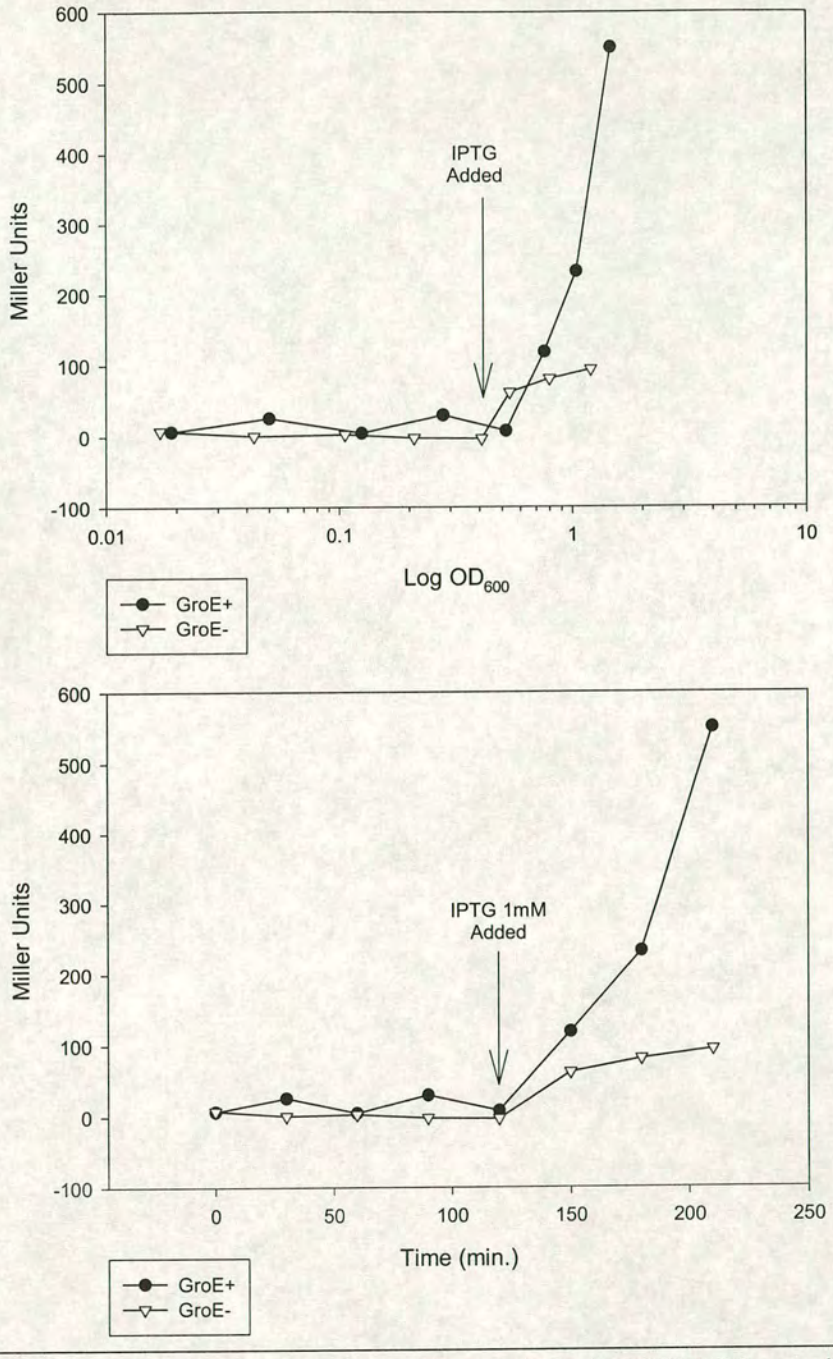


Fig. 4.2.1 Graphs showing β-galactosidase activity in MGM100 growing in buffered LB + Arabinose or Glucose DAP. An overnight culture was diluted 1:100 into buffered LB containing ara, the cells were grown to OD_{600} 0.2, 1 ml of culture was centrifuged, the supernatant was removed and the cell pellet was resuspended in 10 ml buffered LB + ara or glu and DAP (this was time=0). The cultures were induced with 1mM IPTG at 120 minutes. For clarity, when cells have been diluted 1/10, the OD has been multiplied by 10 and plotted on the graph.

4.3 Experiments to test if IPTG enters GroE depleted cells

The gratuitous inducer IPTG can diffuse across the *E. coli* membrane to reach equilibrium with the extracellular concentration. But IPTG is also actively transported across the membrane by the product of the *lacY* gene, *lac* permease, at a rate such that the intracellular to extracellular ratio is 50:1 (Miller, 1972). If LacY were a GroE substrate, and consequently was inactive in GroE depleted cells, then IPTG may be at 50 x higher concentration in $GroE^+$ cells as compared to $GroE^-$ cells. This could explain why LacZ expression is lower in GroE depleted cells compared to GroE normal cells. If this were so, then adding IPTG at an extracellular concentration of 50mM to $GroE^-$ cells should give the same level of LacZ expression as IPTG added at an extracellular concentration of 1mM to $GroE^+$ cells. The results of the β-galactosidase assays are shown in figure 4.3.1. IPTG was added at 1mM and 50mM to both $GroE^+$ and $GroE^-$ cells at 120 minutes. There is no significant difference in LacZ expression in GroE depleted cells at either IPTG concentration. There is also no significant difference in LacZ expression in $GroE^+$ cells at either IPTG concentration, suggesting that 1mM IPTG is enough to cause full induction of the *lac* operon in wild type cells.

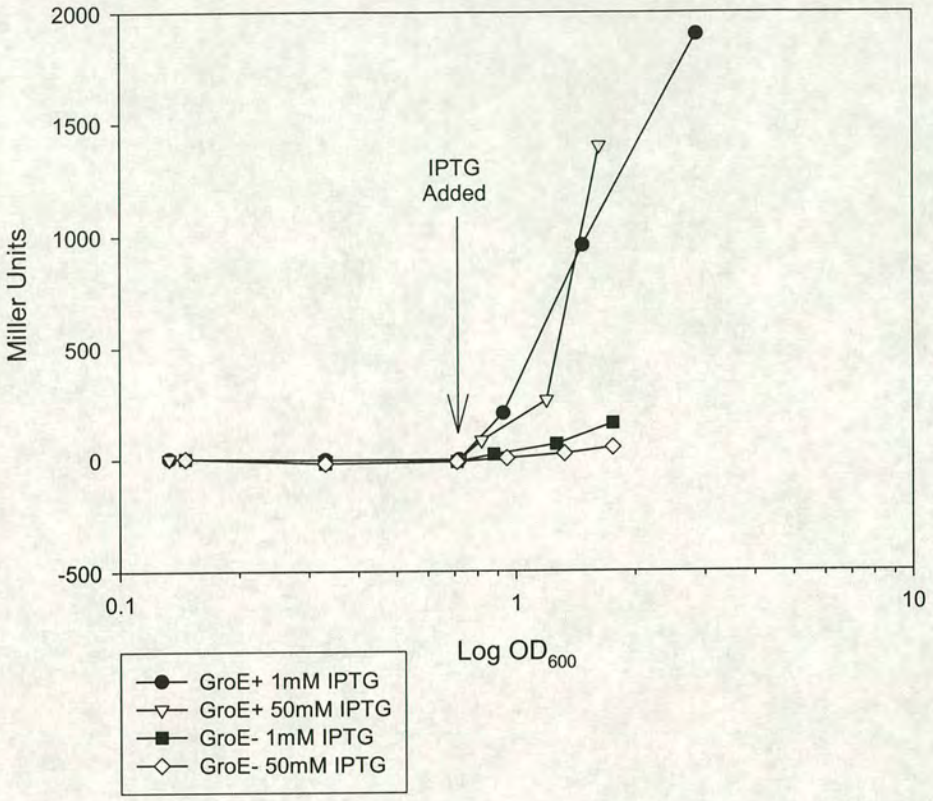


Fig. 4.3.1 β-galactosidase activity in MGM100 growing in buffered LB + Arabinose or Glucose DAP. An overnight culture was diluted 1:100 in to buffered LB containing ara, the cells were grown to OD_{600} 0.2, 1 ml of culture was centrifuged, the supernatant was removed and the cell pellet was resuspended in 10 ml buffered LB + ara or glu and DAP (this was time=0). The cultures were induced with 1mM or 50 mM IPTG at 120 minutes. For clarity, when cells have been diluted 1/10, the OD has been multiplied by 10 and plotted on the graph.

4.4 Construction of MGM100 lacI::Tn10

Since IPTG appears able to enter GroE depleted cells normally, the decreased expression of LacZ upon induction must be due to some other failure in the *lac* regulatory mechanism. As LacI inhibits the transcription of *lac* mRNA, it is possible that LacI was affected by GroE depletion so that it did not dissociate from LacO. To determine whether LacI was involved in causing decreased LacZ expression in GroE depleted cells, a *lacI* mutant of MGM100 was needed. CAG18439, a *lacZ* ochre

mutant is also a *lacI*::Tn*10*. P1 transduction using P1 grown on CAG18439 was used to transduce MG1655 to $LacI^-$, $LacZ^+$. The transduction mix was plated on VB minimal agar plates in the presence of lactose and tetracycline to select $lacZ^+$ and *lacI*::Tn*10* recombinants (*lac* constitutive mutants). One of the MG1655 *lacI*::Tn*10* colonies was then used as the donor to make a new P1 lysate, which was subsequently used to transduce MGM100. Transductants were selected on LB agar in the presence of ara, Kan and tetracycline.

MGM100 *lacI*::Tn*10* has a *lac* promoter that is no longer repressed by LacI; consequently *lacZ* mRNA is constitutively expressed. This strain was grown at 37°C overnight in buffered LB + ara, Kan and tetracycline. The following day the culture was diluted 1/100 into 10 ml fresh buffered LB + ara, and grown at 37°C, with constant shaking until an OD_{600} of ~ 0.2 was reached. 1 ml of culture was then removed, placed into an Eppendorf tube and spun briefly to pellet the cells. The supernatant was removed and the pellet resuspended in 10 ml of either buffered LB + glu, DAP or buffered LB + ara, DAP. The culture was grown for 3 hours with shaking at 37°C, diluting as necessary to keep the OD_{600} below ~ 0.2. Samples were taken every 30 minutes and assayed for β-galactosidase. The results of the assay are shown in figure 4.4.1. In the *lacI*::Tn*10* strain LacZ expression does not decrease as GroE depletes.

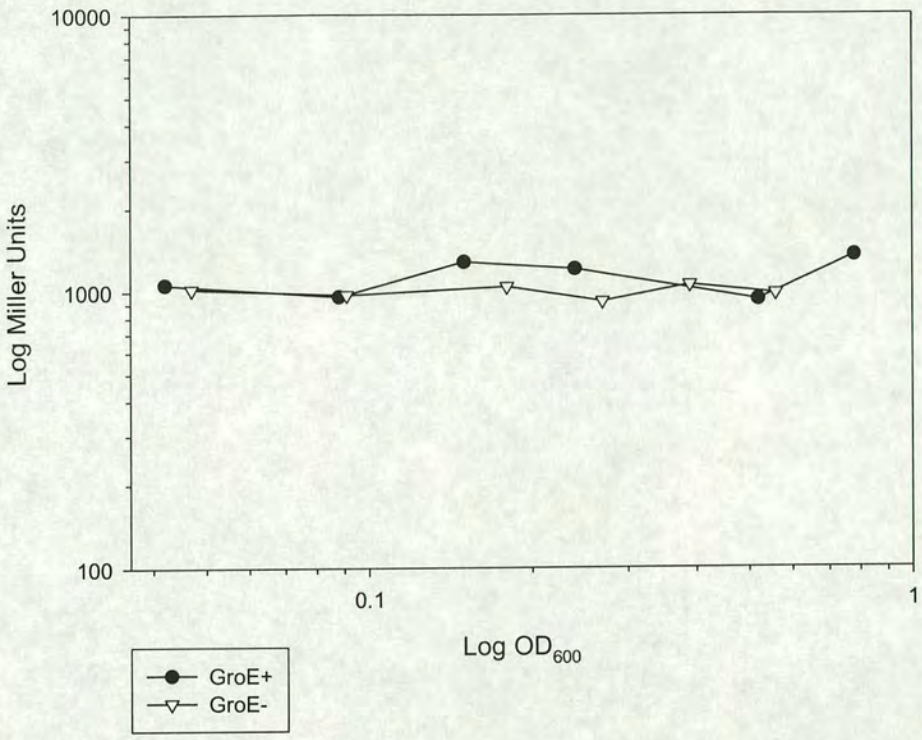


Fig. 4.4.1 β-galactosidase results activity in MGM100 *lacI*::Tn*10* grown in buffered LB + ara, DAP ($GroE^+$) or buffered LB + glucose, DAP ($GroE^-$). An overnight culture was diluted 1:100 in to buffered LB containing ara and Tet, the cells were grown to OD_{600} 0.2, 1 ml of culture was centrifuged, the supernatant was removed and the cell pellet was resuspended in 10 ml buffered LB + ara and Tet or glu, DAP and Tet (this was time=0). Samples were taken every 30 minutes during growth and β-galactosidase activity determined. For clarity, when cells have been diluted 1/10, the OD has been multiplied by 10 and plotted on the graph.

4.5 Constant induction of P_{Lac} with IPTG in GroE depleted cells

The previous experiment had shown that MGM100 *lacI*::Tn*10* did not show a decrease in *lacZ* expression as the intracellular levels of GroE decreased. To determine conclusively that this result was due to the absence of LacI a similar experiment was conducted in a $lacI^+$ strain. By growing MGM100 in the constant presence of 1 mM IPTG, the *lac* operon should always be induced, and then the level of LacZ expression can be measured as the cells are depleted of GroE.

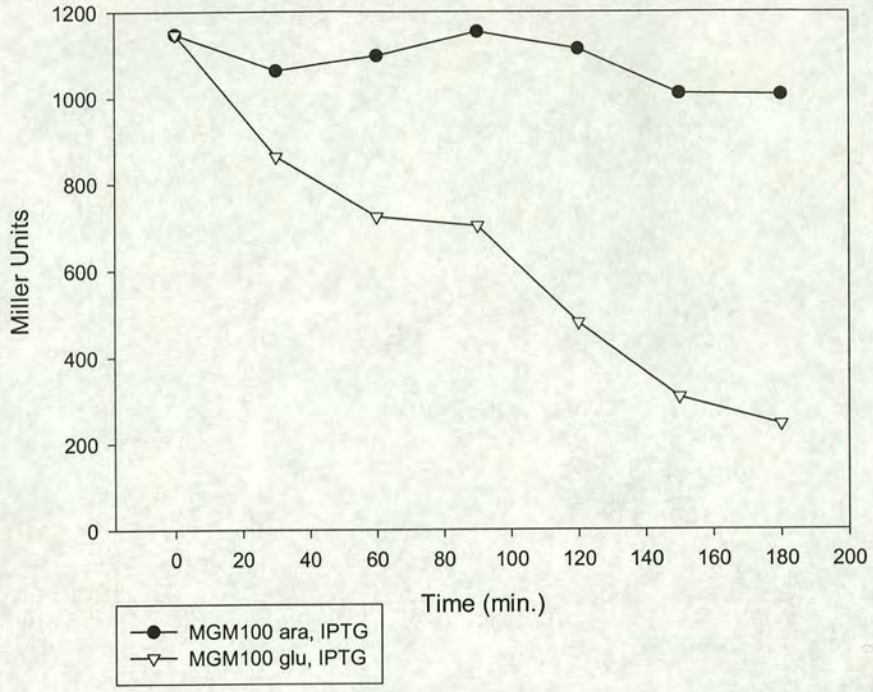


Fig. 4.5.1 β-galactosidase activity in MGM100 grown in buffered LB + glucose, DAP (GroE$^-$) and 1mM IPTG or ara, DAP (GroE$^+$) and 1mM IPTG. An overnight culture was diluted 1:100 in to buffered LB containing ara and IPTG (0.1mM), the cells were grown to OD_{600} 0.2, 1 ml of culture was centrifuged, the supernatant was removed and the cell pellet was resuspended in 10 ml buffered LB + ara and IPTG (0.1mM) or glu, DAP and IPTG (0.1mM).This was time=0. Samples were taken every 30 minutes during growth and β-galactosidase activity determined.

MGM100 was grown as described in section 4.4 except that tetracycline was omitted and 1 mM IPTG was added to the buffered LB (including the overnight culture). The results of this experiment are shown in figure 4.5.1. The β-galactosidase results clearly show that as GroE levels decreased, so did LacZ expression. By comparing the results of this experiment to the result obtained in the previous experiment, it is observed that in a LacI$^+$ strain, as GroE depletes, so does *lacZ* expression; but in a LacI$^-$ strain, *lacZ* expression is unaffected by GroE depletion. Therefore it can be concluded that LacI is affected by GroE depletion such that it remains bound to LacO in the presence of IPTG.

4.6 Footprint analysis of LacI in GroE depleted cells

To prove conclusively that LacI remained bound to LacO when inducer was present it was necessary to 'footprint' the protein DNA complex *in vivo*. The technique of primer extension footprinting analysis (Gralla, 1985) has been used for examining protein-DNA interactions *in vivo* on both plasmid (Sasse-Dwight et al, 1989) and genomic DNA (Sasse-Dwight et al, 1990) in different bacteria (Morett et al, 1988).

The protocol involves three major steps. First, the DNA is covalently modified with an attacking reagent [in this case dimethyl sulfate (DMS)]. Next, the modified DNA is isolated and broken at the sites of modification. Third, the break points are detected by primer extension procedures. In the primer extension a specific ^{32}P end-labelled synthetic oligonucleotide is hybridised to the modified template DNA and extended with a DNA polymerase. The resulting extension products, which all have the same ^{32}P end-labelled 5′ end and a 3′ end that varies according to the position of the template modification at which the primer extension is inhibited, are analysed on a DNA sequencing gel.

If the original modification took place on protein-bound DNA, the attacking reagent may have had altered access to the DNA. In the simplest case this results in a 'protected' band on the gel autoradiograph compared to a lane in which the protein was not present. The interpretation is just as in conventional footprinting techniques that involve using end-labelled template DNA rather than end-labelled primer (Church et al, 1984; Galas et al, 1978).

4.7 Construction of pJF118LacZ

While *in vivo* footprinting can be carried out on either chromosomal or plasmid DNA in *E. coli*, plasmid DNA was chosen for ease of extraction, handling and because the sequence of interest to be footprinted is enriched relative to other sequences. The plasmid pJF118(EH) which contains a *tac* promoter also has $lacI^Q$ on the plasmid, this would provide a higher LacI:LacO ratio than a plasmid binding LacI originating from the chromosome, and should therefore increase the sensitivity of the footprint.

While pJF118(EH) contains a *tac* rather than *lac* promoter, it still contains the LacO operator region and is therefore suitable for use in footprinting the LacI binding site. Also, since the *tac* promoter is a fusion of the –10 region of the *lac*UV5 promoter and the –35 region of the *trp* promoter (Amann et al, 1983), it no longer contained the CAP-cAMP binding site, so this could be excluded from affecting promoter expression in GroE depleted cells. To determine if the *tac* promoter was affected by GroE depletion in a similar manner to the *lac* promoter, *lacZ* was inserted into the multiple cloning site of pJF118(EH). A 4.0kb *lacZ* fragment had been cut from pRS551 using the *Sma*I and *SnaB*I restriction enzyme sites; it had then been ligated into the unique *Sma*I site of pBAD18. The *Sma*I site of pBAD18 is in the multiple cloning site between an upstream *EcoR*I site and a downstream *Hind*III site. To construct pJF118LacZ, the 4kb *lacZ* fragment was cut out of pBAD18 using the *EcoR*I and *Hind*III sites and cloned into the *EcoR*I and *Hind*III sites of pJF118(EH) (Fig 4.7.1.). The ligated DNA was transformed into TP8503, a Δ*lacZ* strain, and colonies were screened on LB plates containing X-gal, IPTG and Amp. Plasmid DNA was extracted from blue colonies, those containing *lacZ* under the P_{tac} promoter, and used to transform MGM100, and the transformed colonies being selected for Kan and Amp resistance.

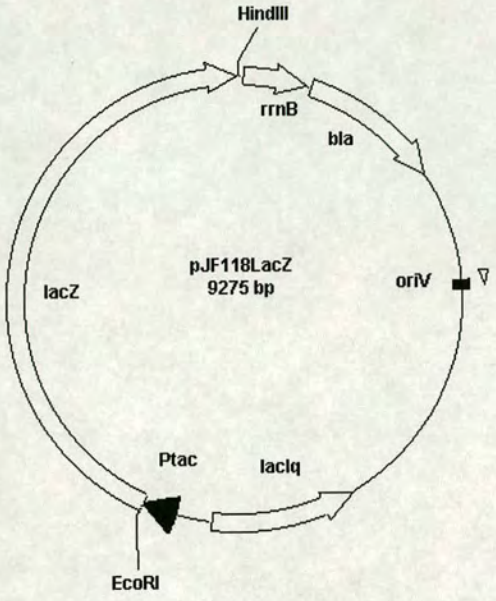


Fig. 4.7.1 pJF118LacZ. The 4kb fragment of *lacZ* from pBADLacZ was cloned into the *EcoR*I and *Hind*III sites of pJF118(EH).

MGM100 pJF118LacZ was grown as per the conditions in section 4.4 except tetracycline was replaced by Amp. Also IPTG, 1mM, was added at 5.5 hours after the initial 1/10 dilution, with samples being taken before and after IPTG addition to use in a β-galactosidase assay. The results of the assay are shown in figure 4.7.2. This data shows that the *tac* promoter behaves similarly to the *lac* promoter when induced in GroE depleted cells, and is therefore suitable for use in footprinting experiments.

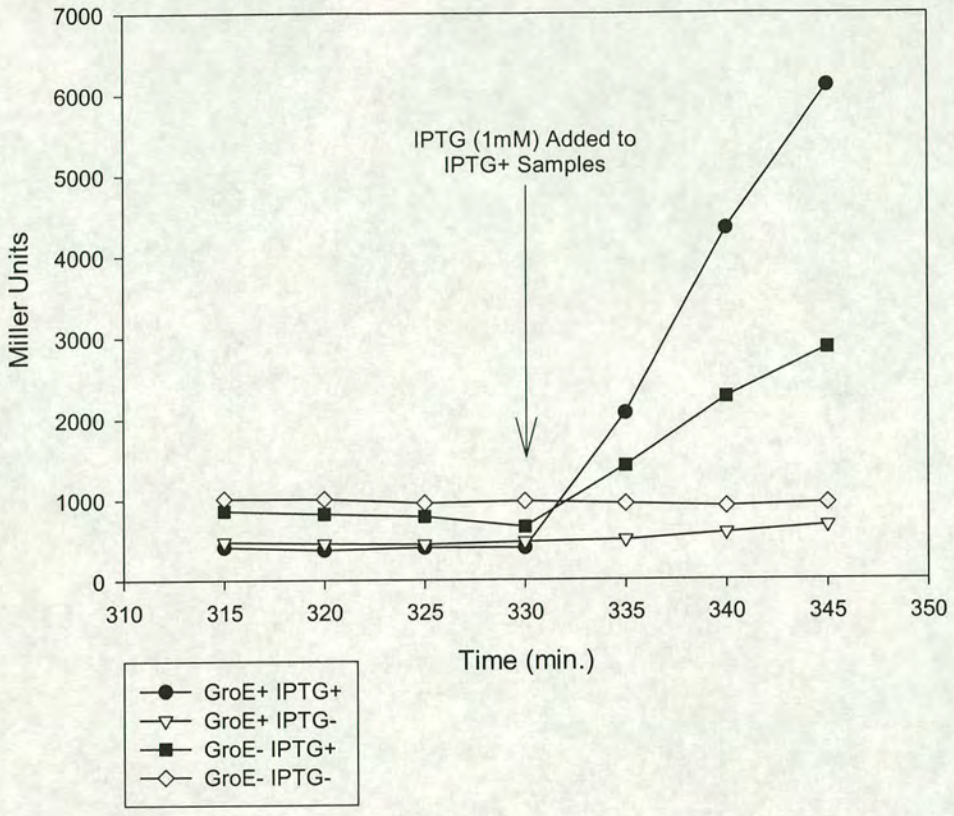


Fig. 4.7.2. β-galactosidase results of MGM100 pJF118LacZ grown in buffered LB + ara, ampicillin, DAP or glucose, ampicillin, DAP. An overnight culture was diluted 1:100 in to buffered LB containing ara and amp, the cells were grown to OD_{600} 0.2, 1 ml of culture was centrifuged, the supernatant was removed and the cell pellet was resuspended in 10 ml buffered LB + ara, amp and DAP or glu, amp and DAP (this was time=0). The IPTG+ cultures were induced with 1mM of IPTG at 5.5 hours.

4.7.1 The lac *operon is not affected by stationary phase growth*

As MGM100 becomes GroE depleted, the generation time of the cells increase; consequently when the *lac* operon is induced with IPTG, the cells at that time have a different growth rate than $GroE^+$ cells (chapter 3.1). While the *lac* operon is not known to be affected by growth rate, there was the possibility that this may be what was causing the observed differences in induction of *lacZ*.

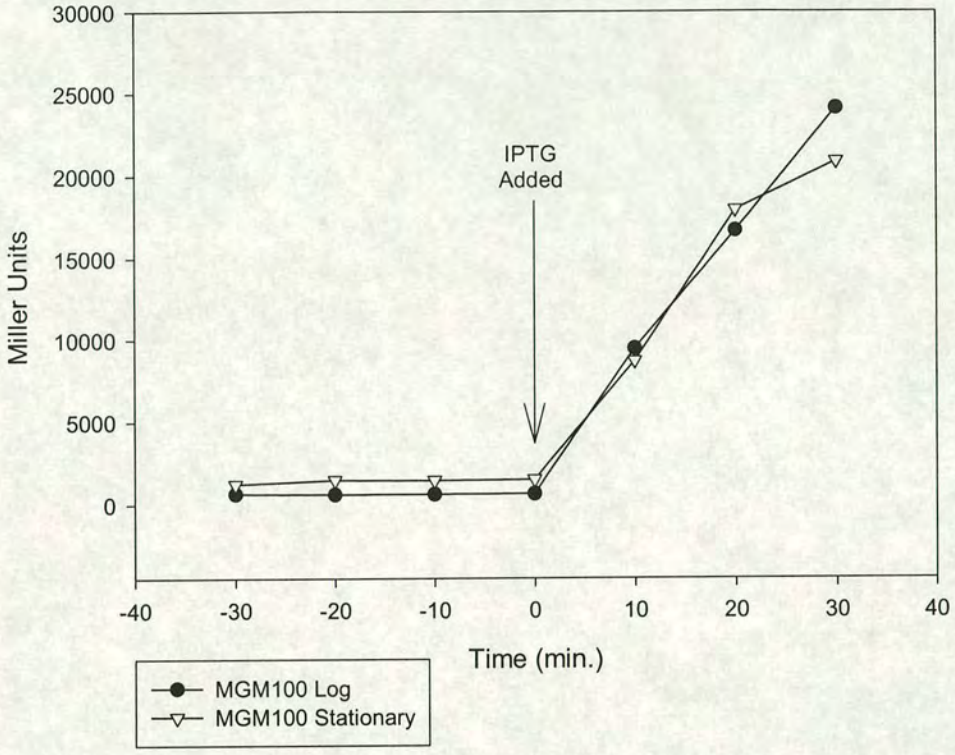


Fig. 4.7.3 MGM100 grown in LB + ara (GroE$^+$) had 1mM IPTG added when the cells were either in logarithmic growth or stationary phase, and the level of LacZ was determined. The X-axis of the graph represents time before and after IPTG addition. Both samples have a similar level of induction, showing that the *lac* operon is unaffected by growth rate.

To determine if the *lac* operon of cells entering stationary phase growth had a different response to IPTG than cells in logarithmic growth, MGM100 was grown in LB + ara (GroE$^+$) to either mid-log (OD_{600} ~ 0.2) or early stationary phase (OD_{600} ~ 1.5) and then 1mM IPTG was added. Samples were taken for 30 minutes before and after the addition of IPTG for use in a β-galactosidase assay. Figure 4.7.3 shows the results of the β-galactosidase assay. These results show that the *lac* operon is unaffected by growth rate, and therefore the decreased levels of expression from the *lac* promoter observed in GroE depleted cells must be caused by GroE depletion.

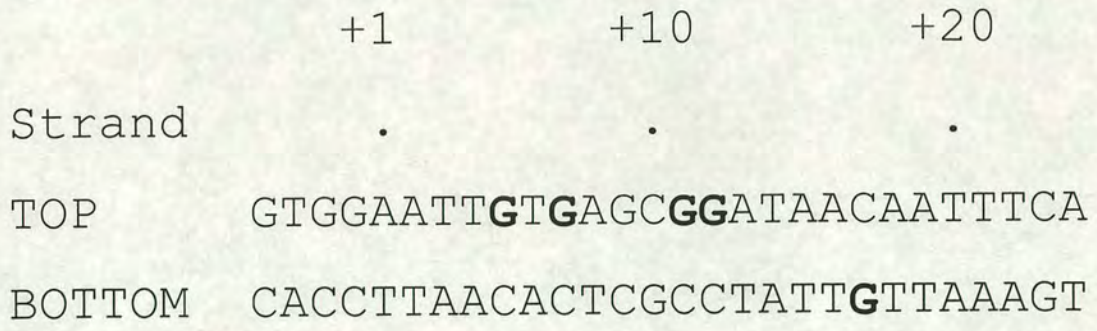


Fig. 4.8.1 LacO DNA sequence with summary of footprint information. The sequence is numbered relative to the position of the transcription start site (+1). Bold lettering delineates DNA methylation protection in the presence of LacI.

4.8 Footprint analysis of LacO

The operator of the *lac* operon was originally identified from cis-acting constitutive (o^c) mutants (Jacob et al, 1961a). A 27 bp DNA fragment, which encompasses these o^c mutants and is protected from deoxyribonuclease digestion by bound *lac* repressor, is centred 11 bp downstream from the start of transcription of the *lacZ* gene (Gilbert et al, 1973). This operator has an axis of approximate dyad symmetry, a feature common to many transcriptional regulatory sites (fig. 4.8.1). Sites within the operator DNA sequence affected by or influencing LacI binding have been identified by a variety of studies using methylation, ethylation, nuclease digestion, photochemical cross-linking, and base substitution (Wick et al, 1991; Barbier et al, 1984; Shalloway et al, 1980). In addition to the principal operator LacO, the *lac* operon has two auxiliary operators O_2 and O_3, which were discovered after the *lac* repressor was isolated and its operator DNA was sequenced (Gilbert, 1975). The centre of the second operator, O_2, is located 401 bp downstream from the primary operator in the *lacZ* gene. The third operator, O_3, overlaps the COOH-terminal coding sequence for the *lacI* gene and is centred 91 bp upstream from LacO (Table 4.8.1). All three operators are required for maximum repression *in vivo* (Oehler et al, 1990).

Operator	Sequence (5′→3′)	Location
LacO	**AATTGTGAGCGGATAACAATT**	+1
O_2	**AAATGTGAGCGAGTAACAACC**	-93 P_{lac}, ~ -250 pJF118
O_3	**GGCAGTGAGCGCAACGCAATT**	+401 P_{lac}, +401 pJF118

Table 4.8.1. Sequences of the *lac* operator, LacO and the pseudo operators O_2 and O_3 aligned with the centre of each operator in the *lac* operon in bold type. Bases listed with capital letters are identical to the bases in the primary operator LacO. Locations are given relative to LacO in either the chromosomal *lac* operon; P_{lac}, or in the plasmid pJF118LacZ; pJF118.

To determine if LacI remained bound to LacO in the presence of inducer in GroE depleted cells, methylation protection footprinting was performed on the plasmid pJF118LacZ in MGM100. A synthetic oligonucleotide, JF118Lac ($^{5'}$GCGCCGACATCATAACGGTT$^{3'}$), was designed which was located 59 bp upstream of the +1 site in the LacO sequence and was complementary to the 'bottom' strand. The oligonucleotide was labelled with [γ-^{32}P] ATP using the T4 polynucleotide kinase reaction as described in materials and methods.

The MGM100 pJF118LacZ samples used to produce the β-galactosidase results in figure 4.7.2 had DMS added to a concentration of 10 mM after the 15 minutes incubation with IPTG. A DMS concentration of 10mM for 3 minutes at 37°C had been show to result in the methylation of less than 1 purine per DNA strand (Ogata et al, 1979). The culture was left with constant agitation at 37°C for 3 minutes, and then transferred to a chilled universal tube and centrifuged for 5 minutes in a chilled rotor at 1500 x g The supernatant was discarded and the DMS modified plasmid DNA extracted according to the protocol given in material and methods.

DMS is capable of penetrating the bacterial cell and modifies the DNA is such a way that the modifications can be made to serve as blocking sites for extension by a DNA copying enzyme. Specifically, DMS modifies, among other sites, the N-7 position of guanine residues. When treated appropriately with piperidine, the DNA strand becomes susceptible to cleavage at these modified guanine residues and this break serves as a block to the DNA copying enzyme during primer extension.

The plasmid DNA was resuspended in 100 µl of 1.0M piperidine and cleaved at 90°C for 30 minutes as described by Maxam and Gilbert (Maxam et al, 1977). The 1 M piperidine was evaporated in a DNA speed vac 110 (Savant) and the cleaved DNA fragments washed twice with 50 µl of water, before being suspended in a final volume of 50 µl of water.

4.9 Primer extension on the P_{Lac} promoter

The primer extension footprinting technique used was a modification of the polymerase chain reaction. This technique amplifies band signal by allowing multiple rounds of extension and permits the extension to occur at a higher temperature, thereby minimising non-specific hybridisation. It is equivalent to PCR except that only one primer is added to the reaction mix. The primer, JF118Lac, since it is complementary to the 'bottom' strand of the plasmid pJF118LacZ would reveal extension termination events in the bottom strand of the LacO operator DNA. The primer extension reaction was performed as per the PCR protocol given in materials and methods. 1.5 µl of the primer JF118Lac (10 pmol μl^{-1}) was added and the reaction was then cycled for 20 rounds of 94°C for 10 seconds, 59°C for 30 seconds and extension temperature of 68°C for 30 seconds, before a final single round of extension for 10 minutes. The PCR mix was then passed through a Sephadex G-25 column (Amersham Pharmacia) and evaporated in a speed vac. The DNA pellet was resuspended vigorously in 5 µl of Sequenase stop solution and 2 µl was loaded onto a 6% sequencing gel (following the protocol in material and methods). The gel was run at 50 Watts until the bromophenol blue dye front reached the bottom of the gel. The gel was then dried and exposed to X-ray film overnight. A manual sequencing reaction of the plasmid pJF118LacZ using unlabelled primer JF118Lac was also run on the same gel so that the appropriate bands could be identified. The results of the primer extension are shown in figure 4.9.1. The lettering down the right hand side of figure 4.9.1 corresponds to the bottom strand of figure 4.8.1 and indicates bases where an extension termination reaction has occurred, i.e. where a base has been modified by DMS. The bold G represents the guanine residue in the bottom strand of the LacO region, which is protected from methylation when LacI is bound. Lanes 1 and 3, corresponding to $GroE^+$ and $GroE^-$ cells without IPTG added, show full protection of the guanine at +17 bp. This protection makes the residue unavailable for modification by DMS, consequently there is no cleavage by piperidine and no extension terminating reaction occurs, so a band is not present on

the gel. This shows that in the absence of IPTG, LacI is bound to LacO in both $GroE^+$ and $GroE^-$ cells. Lane 2 is from $GroE^+$ cells with IPTG added, here LacI is no longer occupying LacO and so the guanine at +17 bp can be modified by DMS, cleaved by piperidine and an extension termination reaction will occur causing a band to appear on the gel. Lane 4 corresponds to $GroE^-$ cells with IPTG added, in this sample there is a small deprotection of the guanine residue at +17 bp, but it is much less than the deprotection we see in the $GroE^+$ sample.

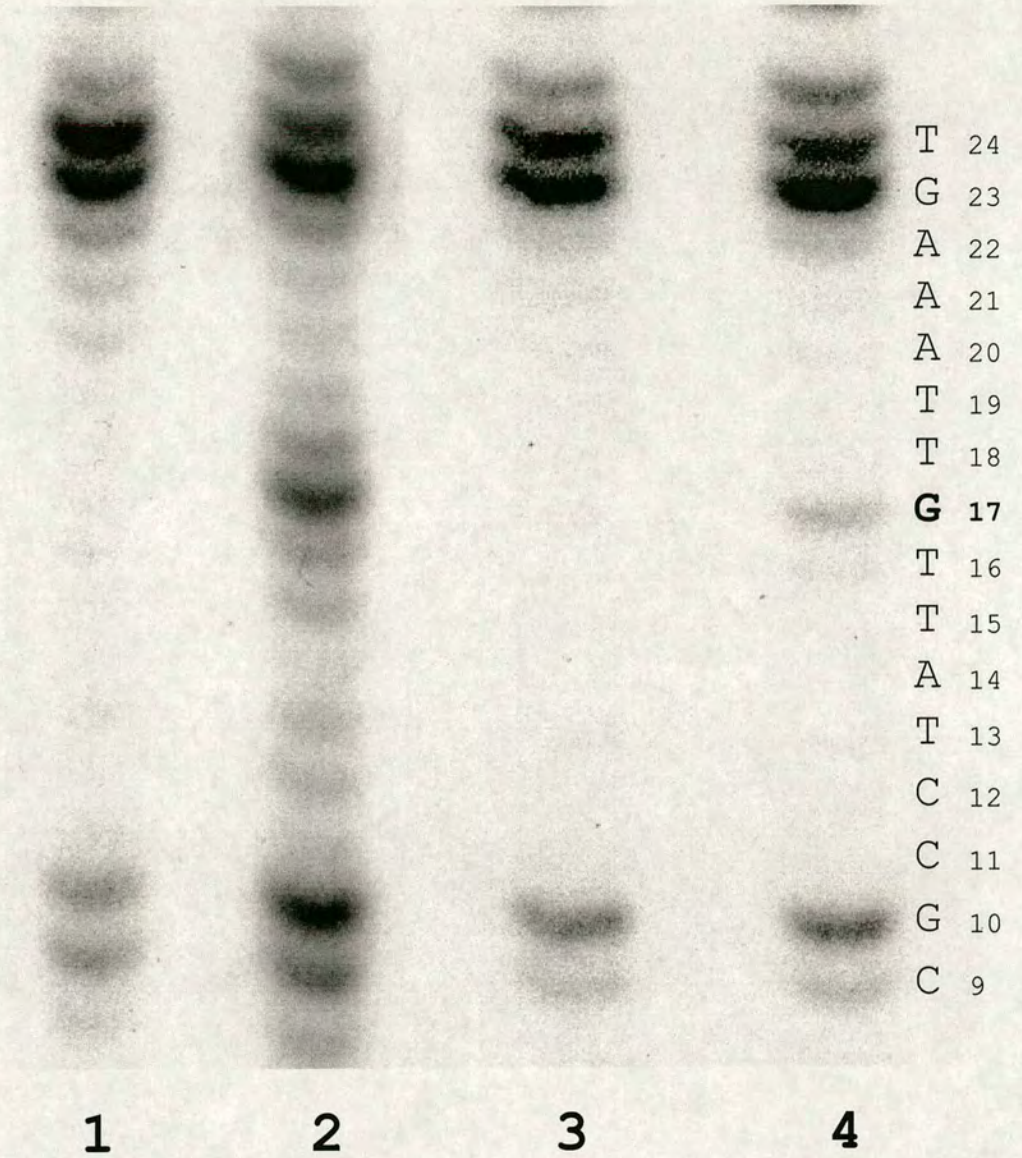


Fig. 4.9.1 Result of the primer extension reaction on pJF118LacZ in MGM100. Lane 1 is $GroE^+$ IPTG-, lane 2 is $GroE^+$ IPTG+, lane 3 is $GroE^-$ IPTG- and lane 4 is $GroE^-$ IPTG+. The lettering down the right hand side of the picture corresponds to the base where the termination reaction has occurred and also corresponds to the bottom strand of figure 4.8.1.

To resolve any errors that may have been caused by loading differences on the gel or different amounts of template DNA used in the primer extension reaction, a control band was chosen, the DMS modification of which is not affected by LacI, in this case the guanine at +23 bp, and the ratio of the intensity of guanine 17 to guanine 23 determined for both the $GroE^+$ and $GroE^-$ samples (the bases G10 and T24 are both affected by LacI binding (Matthews et al, 1998) and were therefore not suitable for use as control bands).

The results of this area integration are shown in figure 4.9.2. The baseline for these graphs can be determined in one of two ways, using either a baseline of best fit (solid line) or a lowest value baseline (dotted line). Though using either baseline gives approximately the same peak-to-peak ratio. For the purposes of this analysis, the lowest point baseline was used. For the $GroE^+$ sample with IPTG present, the area of peak G 23 is 4928 (arbitrary units) and the area of peak G 17 is 3619, giving a ratio of G17:G23 of 0.73. For The $GroE^-$ sample with IPTG present, the area of peak G 23 is 4819 and the area of peak G 17 is 856, giving a G17:G23 ratio of 0.18. Therefore the percentage deprotection of the guanine 17 for the $GroE^-$ compared to the $GroE^+$ sample is 24.7%. (Using a best-fit baseline this value is 23%). From this data it can be concluded that in a $GroE^-$ cell, in the presence of 1mM IPTG, only ~ 25% of LacI has dissociated from LacO when compared to a $GroE^+$ cell.

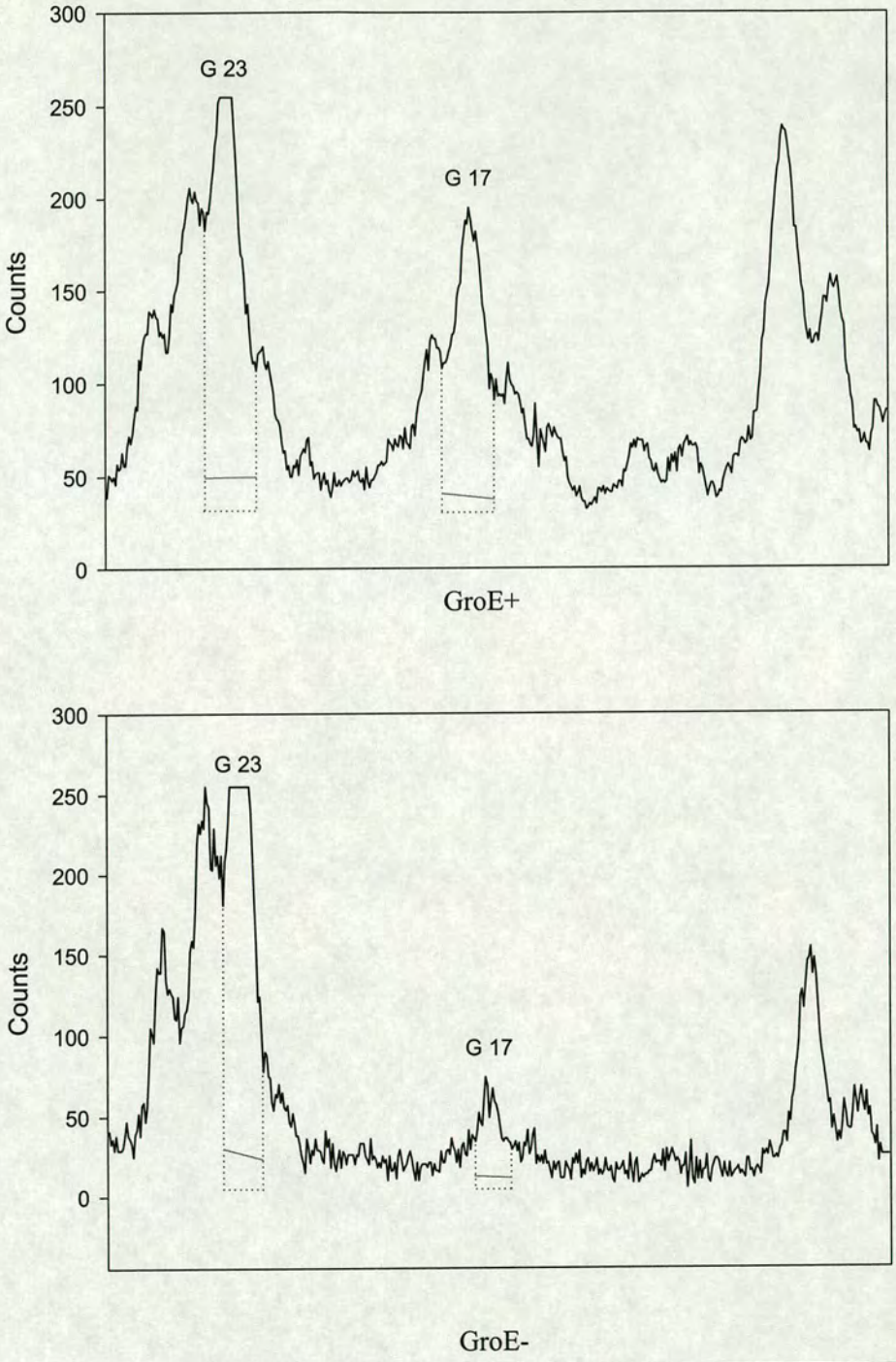


Fig. 4.9.2 Intensities of bands in figure 4.4.4 in $GroE^{+}$ and $GroE^{-}$ samples incubated with 1mM IPTG for 15 minutes. G 23 and G 17 represent the intensity of the bands on figure 4.4.4 representing guanines at 23 bp and 17 bp respectively

4.10 Summary and Discussion

Initial experiments in the study of the *lac* promoter showed that *lacZ* was not expressed at the same level in GroE depleted cells as in $GroE^+$ cells in the presence of the gratuitous inducer IPTG. Previous studies in this work had eliminated LacZ folding or function from being the cause of this problem as LacZ can be expressed from other, non-*lac* promoters in GroE depleted cells to the same level as in $GroE^+$ cells. Therefore, the decreased level of expression of *lacZ* must have resulted from a decreased level of activity in the *lac* promoter.

lac permease, encoded by the *lacY* gene is responsible for the transport of lactose across the *E. coli* membrane. IPTG can diffuse across the bacterial membrane to reach equilibrium with the external concentration, but LacY can also transport IPTG such that the internal concentration is 50 times higher than the external concentration. If LacY were affected by GroE depletion, for example not folding correctly and therefore being unable to transport IPTG across the membrane, then there may have been a reduced internal concentration of IPTG in $GroE^-$ cells. If LacY was not transporting IPTG across the membrane then increasing the extracellular IPTG concentration to 50mM should have produced the same level of *lacZ* expression in a $GroE^-$ cell as 1mM IPTG would produce in a $GroE^+$ cell. Adding IPTG at 50 fold higher concentration did not affect *lacZ* expression in GroE depleted cells, and from this it was concluded that LacY did not seem to be the factor limiting expression. This could also be tested by measuring the uptake of radioactively labelled IPTG in a $GroE^{+/-}$ strain.

The other obvious protein, which might affect expression from the lac promoter, is the *lac* repressor protein: LacI. This protein binds to the *lac* operator sequence, LacO, preventing RNA polymerase from binding to the DNA and transcribing *lac* mRNA. In the presence of either β-1,6-allolactose or IPTG, LacI undergoes a conformational change that causes it to have decreased affinity for LacO, such that RNA polymerase can now bind and transcription of *lac* mRNA proceed. If LacI were affected by GroE depletion so that it did not have reduced affinity for LacO in the presence of IPTG, this might account for the reduced level of expression from a *lac* based promoter in GroE depleted cells.

In MGM100 a *lacI*::Tn*10* strain was constructed. This strain, regardless of the amount of GroE, constitutively expressed *lacZ* mRNA. In contrast LacZ expression decreased markedly in MGM100 ($lacI^+$) depleted of GroE in the continual presence of IPTG. From these two experiments it was possible to deduce that LacI was being

affected by GroE depletion and that it did not have reduced affinity for LacO when the inducer IPTG was present.

To determine conclusively that LacI remained bound to LacO in the presence of inducer, methylation protection footprinting was performed on the LacI-LacO complex. In this procedure a DNA methylating agent, DMS, is added to $GroE^{+/-}$ cells in the presence/absence of IPTG. If LacI is bound to LacO it will protect certain bases from methylation that would otherwise be methylated, these bases having been well characterised in previous works. The methylating agent is added at a concentration such that less than 1 purine per DNA molecule is modified. The modified DNA is then cleaved at the methylated bases and used as the template in a primer extension reaction. The resulting footprint, seen on the sequencing gel, represents bases that have been modified by DMS. If LacI was bound to LacO then certain bases should have been protected from methylation and should be absent when compared to a sample where LacI is disassociated from LacO. From the footprint we can clearly see that in a GroE depleted cell greater than 75% of LacI remains bound to LacO in the presence of IPTG.

The main reason that DMS was used as the DNA modification reagent in this work is that the interaction between LacI and LacO had been previously characterised using DMS footprinting, and thus the data on bases protected by LacI was readily available. The use of other DNA modification reagents to give a more detailed picture of bases affected by LacI binding should not be ignored. For example, DMS modifies the N-7 position of guanine residues and potassium permanganate ($KMnO_4$) modifies primarily thymine and to a lesser extent cytosine residues (Hayatsu et al, 1967). In addition to these two chemicals, several enzymes have been used in primer extension footprinting. These include DNaseI (Buchanan et al, 1987; Zhang et al, 1990) and micrococcal nuclease (Zhang et al, 1989), which have different modes of cleavage and thus give complimentary information. They are not suitable for use in bacteria since they do not penetrate the cells, but have been used in primer extension protocols *in vitro* and in permeabilised mammalian cells (Zhang et al, 1989).

A major limitation in this work has been the inability to perform the primer extension reaction on the top stand of LacO. While one base on the bottom strand of LacO is affected by the presence or absence of LacI, the top strand would have been preferable to use as the methylation states of 4 bases are affected (Fig. 4.8.1). Unfortunately the primer extension reaction failed to work on the top strand. Multiple primers were used with a variety of different binding sites and different T_m's, different polymerases were used as well as different primer extension

conditions, but none of these were able to overcome the problem. The primer used in this work, JF118Lac, was complementary to a region of DNA upstream of LacO, but the primers complementary to the top strand of LacO bound downstream of it. In the plasmid pJF118LacZ there is a multiple cloning site just downstream of LacO and since most restriction enzyme sites display twofold symmetry it is possible that areas of secondary structure may have formed (DNA hairpins) in the multiple cloning site. If this were the case then it could explain why primer extension could not proceed on the top strand; the polymerase, upon reaching the area of secondary structure, would stall and extension would terminate. This hypothesis is consistent with the results observed upon probing the top strand, the sequencing gel showing primer extension products a short distance away from the primer, but then the absence of any more bands before the LacO site is reached.

Further work on this subject would be to determine how LacI is affected by GroE depletion, whether it is an indirect effect, i.e. is DNA structure altered in GroE depleted cells causing the 'off' rate of LacI in the presence of IPTG to decrease, or is it a more direct effect, such as misfolding of LacI protein.

It is interesting to note that when LacI was overproduced (plasmid pJF118LacZ also has $lacI^{Q}$ on it, this carries a mutation in the *lacI* promoter region which causes an increased level of expression of *lacI*) the expression of LacZ was not affected until much later into GroE depletion than when LacI was present in single copy on the chromosome. In figure 4.2.1 IPTG is added to cells containing the chromosomal *lac* promoter at 120 minutes and there is a noticeable difference between LacZ expression in $GroE^{+}$ and $GroE^{-}$ cells. When this experiment is performed using the plasmid pJF118LacZ (Fig. 4.7.2) the IPTG is added at 330 minutes to observe the same result, adding IPTG much before this time results in similar levels of LacZ expression in both $GroE^{+}$ and $GroE^{-}$ cells. There seem to be some parallels with the overproduction of LacI prolonging the LacI response to IPTG in GroE depleted cells and of DapA delaying lysis in GroE depleted cells (McLennan et al, 1998). It may be possible that, like DapA, as an excess of LacI is made when GroE is still present, it takes longer to reduce this intracellular pool of LacI which has not been affected by GroE depletion, consequently the *tac* promoter would not be affected by GroE depletion until much later.

The LacI monomer has four functional units, an NH_2-terminal headpiece which binds specifically to the DNA, the hinge region, a sugar binding domain and a COOH-terminal helix, which is responsible for LacI tetramer formation (Lewis et al, 1996). LacI binds to LacO as a dimer, with each NH_2-terminal headpiece binding to one half of the palindromic LacO sequence, yet the fully functional *lac* repressor is a

tetramer. To achieve maximal repression of the *lac* operon requires three operators suggesting that the tetrameric form of the repressor is required for maintaining the physiology of the system. Removing LacO, O_2 or O_3 (Table 4.8.1) reduces the repression ratios by varying degrees (Oehler et al, 1990). A single *lac* repressor tetramer binds two operators that are separated by 401 bp or 93 bp in a *lac* promoter (401 bp and ~ 250 bp in the plasmid pJF118 LacZ). Since the repressor tetramer binds two independent pieces of operator DNA, a continuous DNA must bend to form repression loops (Eismann et al, 1990; Flashner et al, 1988) (Fig. 4.10.1).

The results observed when *lac* or *tac* promoters are induced in GroE depleted cells do not show a decreased level of repression, but rather an increased level of repression. This suggests LacI is probably still present as a tetramer, is still forming a repression loop and that any possible misfolding is not affecting the COOH-terminal helix.

DNA supercoiling is affected in *groE* strains (Murayama et al, 1996). It has also been reported that expression from the *lac* promoter is decreased in strains where DNA gyrase is inactive (Yung et al, 1979; Borowiec et al, 1979). While this may be a possible explanation for the decreased expression from P_{Lac} in a GroE depleted cell it seems unlikely. DNA supercoiling affects expression from P_{Lac} by decreasing the affinity between RNA polymerase and the *lac* promoter, not by increasing the affinity of LacI for the operator region. Also while changes in supercoiling may render certain bases in the *lac* promoter region more or less susceptible to modification by DMS, the base G17 has not been reported to be affected by supercoiling. This suggests that the change in methylation sensitivity of G17 is due solely to the presence or absence of LacI and thereby further suggests that the presence of LacI and not the absence of RNA polymerase is responsible for the decreased expression from P_{Lac} in a GroE depleted cell plus IPTG.

The three most plausible areas of LacI which, if misfolded could cause the protein to become unresponsive to IPTG would be, (i) the DNA binding domain; a mutation in this domain might cause LacI to remain irreversibly bound to LacO, (ii) the sugar binding domain; a mutation in this domain could cause LacI to not interact with IPTG and consequently no conformational change would come about to cause LacI to have reduced affinity for LacO or (iii) the hinge region; even if the sugar binding domain and the DNA binding domain were functional, without the hinge region there would be no 'communication' between either domain and thus LacI may not respond to the presence of IPTG.

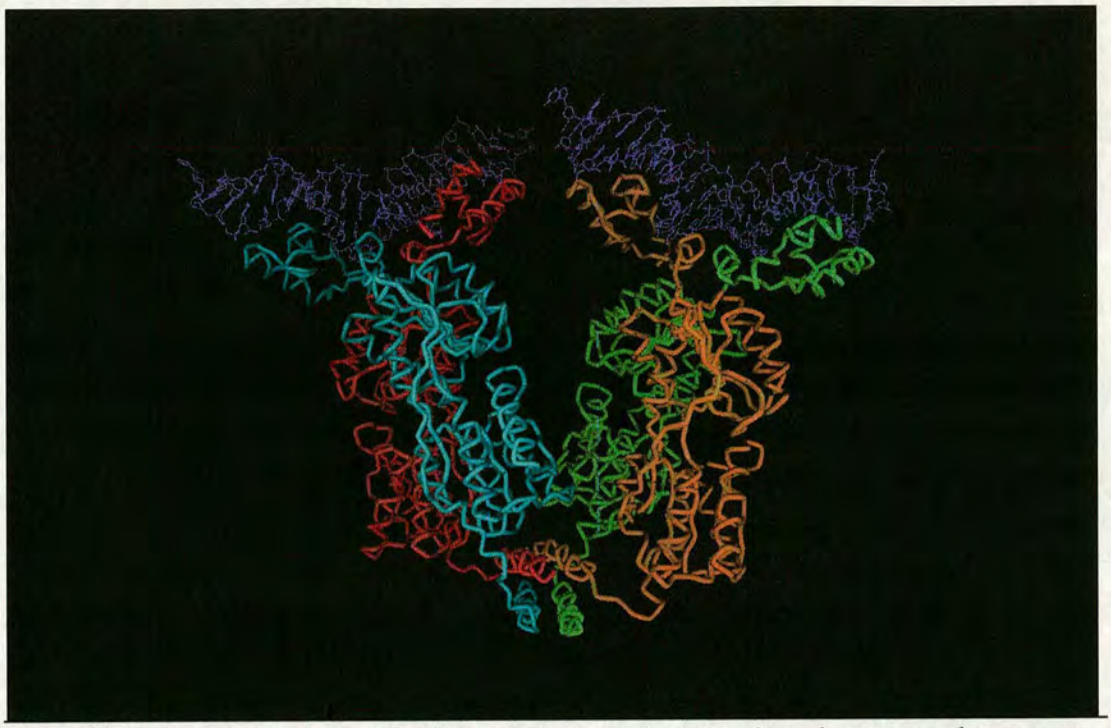


Fig. 4.5.1. Representation of LacI tetramer bound to LacO and one pseudo operator. LacI monomers are coloured red, light blue, green and bronze, with the COOH-terminal helices, responsible for tetramer formation at the bottom of the picture.

To determine if the LacI DNA binding domain were misfolded in GroE depleted cells causing it to remain irreversibly bound to LacO, DNA footprinting could be performed. If LacI was bound to the operator DNA in an altered conformation then it may also exhibit an altered DNA footprint. To confirm this a number of reagents would have to be used to probe the DNA. One advantage to using a variety of reagents is that each attacks the DNA differently and their joint use can give a very detailed picture of the DNA bound by the protein. For example, DNaseI cleavage occurs by attack through the DNA minor groove, DMS attacks guanines in the DNA major groove, micrococcal nuclease attacks exposed positions on the DNA backbone, and permanganate attacks melted DNA preferentially (Borowiec et al, 1987; Zhang et al, 1990; Zhang et al, 1989). Therefore by comparing the footprint of LacO from $GroE^-$ compared to $GroE^+$ cells it may be possible to determine if the LacI DNA binding domain is affected by GroE depletion.

In conclusion, the isolation of LacI from GroE depleted cells and the determination of its structure may extend our knowledge of potential GroE substrates *in vivo*, and further our understanding of the mechanisms involved in the regulation of the *lac* operon.

Chapter V

Regulation of *dapA* expression

5.1 Introduction

The *dapA* gene encodes dihydropicolinate synthetase (DHDPS, E.C. 4.2.1.52), the first enzyme of the diaminopimelate and lysine pathway (Shedlarsky et al, 1970). It has been mapped to 55.8 minutes on the *E. coli* chromosome (Bachmann, 1983). DHDPS activity is sensitive to lysine inhibition (Yugari et al, 1965) as is aspartokinase III, encoded by the *lysC* gene (Stadtman et al, 1961), the first enzyme on the biosynthetic pathway leading from aspartate to diaminopimelate, lysine, methionine, threonine and isoleucine (Fig 5.1.1). The regulation of expression of several genes of the lysine regulon (Patte, 1983), including *lysC* (Stadtman et al, 1961), *asd* (Boy et al, 1972), *dapB* (Losick et al, 1966), *dapD* (Richaud et al, 1984), *dapE* (Cohen et al, 1965), and *lysA* (Patte et al, 1962), has been studied either by directly measuring their gene products or by the use of *lacZ* or *galK* fusions. In contrast to all these genes, whose level of expression respond to lysine pool size, the synthesis of DHDPS has not been reported to be subject to any regulatory control (Butour et al, 1974; Bouvier et al, 1984).

The native DHDPS enzyme forms a homotetramer with a molecular weight of 31,372 Da per monomer (Richaud et al, 1986; Shedlarsky et al, 1970). The amino acid sequence has been derived both from the nucleotide sequence and protein analysis (Laber et al, 1992). Also amino acid sequences of DHDPS of *Corynebacterium glutamicum* (Bonnassie et al, 1990), wheat (Kaneko et al, 1990) and maize (Frisch et al, 1991) have been published. In all of these cases, DHDPS is a homotetramer as well. In plants, the enzyme is competitively inhibited by low concentrations of the endproduct of the pathway, lysine, with $I_{0.5}$ values ranging from 0.01 to 0.05 mM (Perl et al, 1992; Shaul et al, 1993; Cheshire et al, 1975; Ghislain et al, 1990; Frisch et al, 1991; Dereppe et al, 1992; Kumpaisal et al, 1987). Bacterial DHDPS is inhibited to a much lesser extent with an $I_{0.5}$ value of 1.0 mM (Laber et al, 1992).

In this chapter, work is described which suggests that the gene *dapA* is regulated at the level of transcription or translation and that the expression of *dapA* is inversely proportional to the intracellular levels of the compound meso-diaminopimelate.

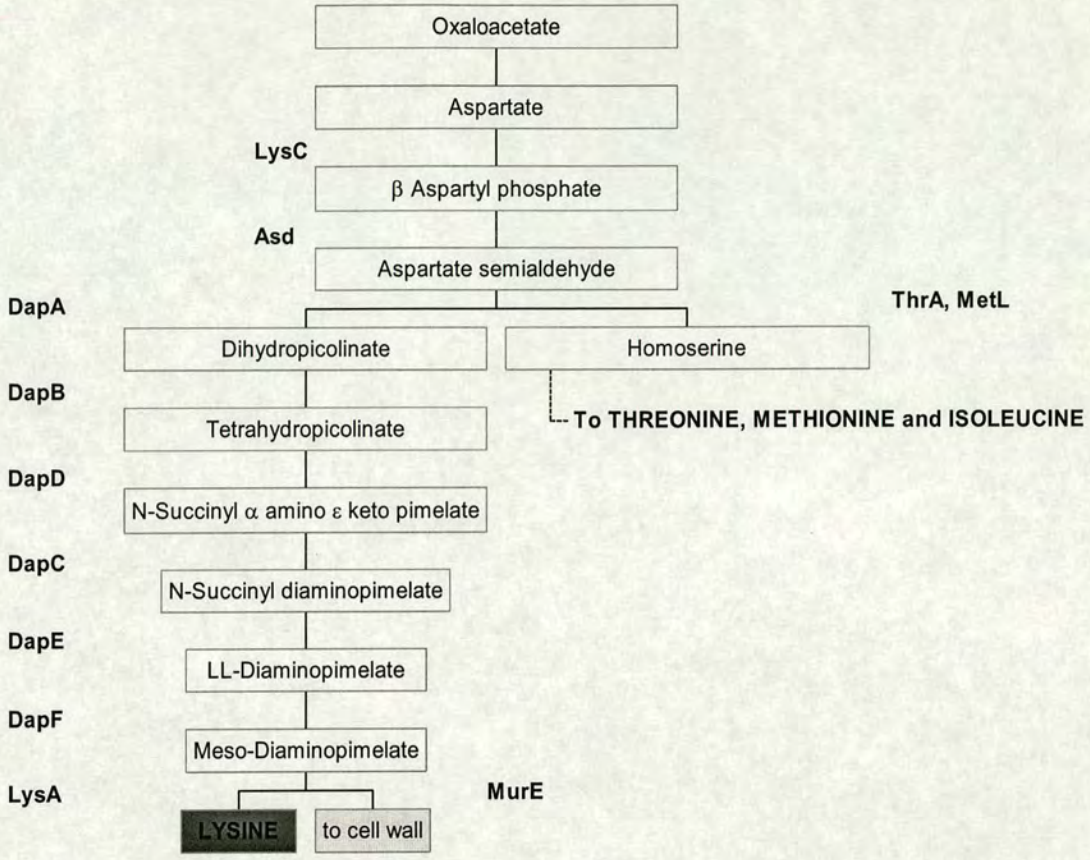


Fig. 5.1.1 The biosynthetic pathway from oxaloacetate to lysine, threonine, methionine and isoleucine. The product of the *lysC* gene is aspartokinase III; *asd*: aspartate semialdehyde dehydrogenase; *dapA*: dihydropicolinate synthetase; *dapB*: dihydropicolinate reductase; *dapD*: tetrahydropicolinate succinylase; *dapC*: succinyl diaminopimelate aminotransferase; *dapE*: succinyl diaminopimelate desuccinylase; *dapF* diaminopimelate epimerase; *lysA*: diaminopimelate decarboxylase; *murE*: UDP-N-acetyl-muramoylalnyl-D-glutamate-2,6-diaminopimelate ligase; *thrA*: homoserine dehydrogenase I; *metL*: homoserine dehydrogenase II.

5.2 Increased *dapA* expression in GroE depleted *E. coli*

Using the strain MGM100λP_{dapA}-*lacZ* it was previously observed that, just before lysis caused by GroE depletion, expression of LacZ under the control of the *dapA* promoter increased in cells growing in LB (N. McLennan, pers. comm.). The level of expression of LacZ from the *dapA* promoter in a $GroE^+$ strain is ~ 300 Miller units, while from a $GroE^-$ strain just before lysis it is ~ 650 Miller units. This suggests that, contrary to previous observation, *dapA* is regulated at the level of expression and that there is increased transcription or translation of DapA in GroE depleted cells.

Since it was known that in GroE depleted cells, the enzyme DapA had little activity, the first step was to determine if *dapA* expression was autoregulated.

A P1 lysate was made using MGM100Δ*dapA* as the donor. This lysate was then used to transduce MGM100λP_{dapA}-*lacZ*. Transductants were selected on minimal media plates + Cmp and DAP. Those colonies which were Kan^R, Cmp^R and required both ara and DAP for growth on minimal medium (MGM100Δ*dapA*λP_{dapA}-*lacZ*) were then re-streaked on suitable plates or frozen for later use.

The strain MGM100Δ*dapA*λP_{dapA}-*lacZ* was grown in minimal medium, casamino acids (CAA), ara and DAP; these conditions resulted in a $GroE^+$, $dapA^-$ strain. MGM100λP_{dapA}-*lacZ* was grown both under these conditions and also in the absence of DAP. The strains were grown overnight at 37°C with constant agitation and then diluted 1:100 into fresh medium the next day. The cells were grown in a 37°C water bath with shaking until the OD_{600} reached 0.2, samples were then taken and used to determine the β-galactosidase activity. The results of five repeats of this experiment are shown in Table 5.2.1.

The results show that in all cases β-galactosidase activity and hence *dapA* promoter activity remains approximately the same. This would indicate that *dapA* expression is not autoregulatory, as a $dapA^-$ strain has the same level of expression from P_{dapA} as a $dapA^+$ strain.

In the following experiments, MG1655 could have been used instead of MGM100, but since the initial observations regarding *dapA* expression were made in MGM100, this strain was used.

Strain	*DapA* genotype	DAP added to medium	Miller Units #1	Miller Units #2	Miller Units #3	Miller Units #4	Miller Units #5	Average Miller Units
MGM100Δ*dapA* λP_{dapA}-*lacZ*	*dapA*$^-$	Yes	388	361	379	365	378	374.2
MGM100 λP_{dapA}-*lacZ*	*dapA*$^+$	Yes	385	345	360	366	379	367
MGM100 λP_{dapA}-*lacZ*	*dapA*$^+$	No	382	424	386	404	388	396.8

Table 5.2.1 Miller units representing the level of β-galactosidase activity from λP_{dapA}-*lacZ* in either a *dapA*$^+$ or *dapA*$^-$ strain of MGM100 grown under various conditions. The strains were grown in minimal medium until OD_{600} ~ 0.2 and samples were taken to determine β-galactosidase activity. The experiment was repeated 5 times. Since all samples have a similar level of expression from P_{dapA} this suggests that *dapA* expression is not autoregulatory.

5.3 *dapA* expression is regulated by a downstream product

The data above shows that *dapA* is not autoregulated, but it was possible that *dapA* expression may be regulated by a downstream product. Since lysine levels were known not to affect *dapA* expression (Bouvier et al, 1984) the next compound to check was meso-diaminopimelate (DAP). DAP is at the branch-point of the lysine biosynthetic pathway which leads either to lysine or to incorporation into the cell wall by MurE, and therefore would be a possible regulatory compound. In the previous experiments all the cells contained DAP, either from an extracellular source or because they could synthesise it (MGM100Δ*dapA*λP_{dapA}-*lacZ* requires the presence of DAP to grow and MGM100λP_{dapA}-*lacZ* is a *dapA*$^+$ strain and therefore will produce DAP). Therefore if DAP does regulate *dapA* expression it is not surprising that the previous samples had a similar level of expression from P_{dapA}.

To test this hypothesis it was necessary to grow a *dapA*$^-$ mutant without DAP. This is problematic as cells grown under these conditions will lyse as they cannot make peptidoglycan (the result observed in GroE depleted cells). To overcome this

problem the cells were grown in osmotically protected medium (medium containing 10 mM $MgSO_4$), which prevents the cells from lysing (Heijenoort, 1996).

The strain MGM100Δ*dapA*λP_{dapA}-*lacZ* was grown as described in section 5.2, except that when the cells reached an OD_{600} of 0.2, 1 ml of culture was removed, centrifuged and the supernatant removed. The cell pellet was then resuspended in 10 ml of minimal medium, CAA, with 10 mM $MgSO_4$, ara and DAP or just 10 mM $MgSO_4$ and ara (this was time=0). Samples were then taken every 10 minutes and β-galactosidase activity determined.

The results (Fig. 5.3.1) show that expression of LacZ from P_{dapA} increases in the strain MGM100Δ*dapA*λP_{dapA}-*lacZ* when extracellular DAP is removed, while expression from P_{dapA} remains unaffected when extracellular DAP is present. These results were explained by the hypothesis that DAP regulates expression from the promoter P_{dapA}. In MGM100Δ*dapA*λP_{dapA}-*lacZ*, intracellular DAP levels decrease when DAP is removed from the medium. As the strain is *dapA*$^-$ and cannot synthesise DAP, the intracellular 'pool' of DAP is exhausted as it is incorporated into the peptidoglycan cell wall; the decreasing levels of DAP therefore lead to an increase in the level of expression from the promoter P_{dapA}. If DAP is supplied in the medium it is transported inside the cell and therefore intracellular DAP levels do not decrease and subsequently expression from P_{dapA} is unaffected.

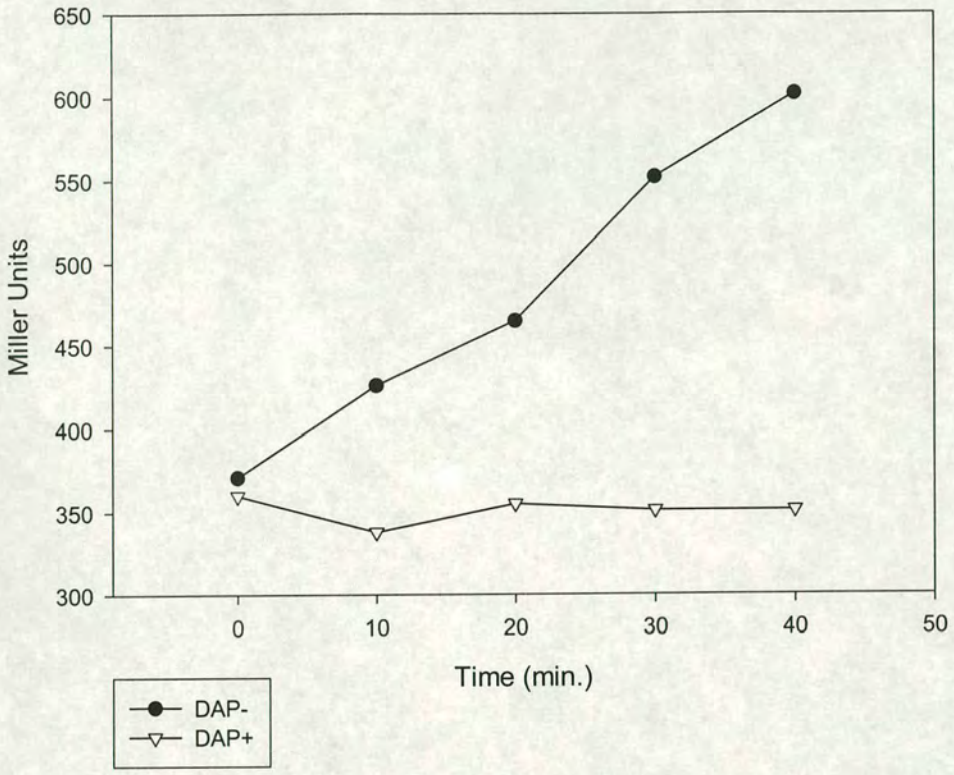


Fig. 5.3.1 β-galactosidase activity in MGM100Δ*dapA*λP_{dapA}-*lacZ* growing in osmotically protected minimal medium, CAA, ara with (DAP+) or without (DAP-) DAP. The strain was grown in minimal medium, CAA, ara and DAP and then centrifuged, the supernatant removed and the cells resuspended in osmotically protected minimal medium + ara and DAP or just ara (this was time=0). As the intracellular levels of DAP fall in the strain grown without DAP, expression from P_{dapA} increases.

5.4 Lysine levels also affect expression from P_{dapA}

The previous data suggests that the intracellular levels of the lysine precursor DAP regulates expression from P_{dapA}. The reason therefore that expression from P_{dapA} increases in GroE depleted cells was that as DapA was not functional (see chapter 3), DAP was not synthesised, and consequently intracellular levels of free DAP decreased as it was incorporated into peptidoglycan (if DAP was supplied in

the medium, increased expression from P_{dapA} was not observed during GroE depletion).

In all the experiments up until this point lysine has also been present in the medium, and therefore intracellular DAP was only needed for peptidoglycan synthesis. If lysine were missing from the medium, then intracellular DAP levels would be reduced more quickly as it would be both incorporated into peptidoglycan and used to synthesise lysine.

The strain MGM100Δ*dapA*λP_{dapA}-*lacZ* was grown as described in section 5.3, except that instead of minimal medium and CAA, the cells were grown in minimal medium and SCAA –Lys (see chapter 6). This is a synthetic medium that contains the same levels of each amino acid as would be supplied using casamino acids except that it has no lysine.

When the cells were grown in this medium in the absence of DAP, there was a great increase in the level of LacZ, thereby showing a great increase in the level of expression from P_{dapA} (Fig 5.4.1). In the graph, β-galactosidase results from MGM100Δ*dapA*λP_{dapA}-*lacZ* grown in osmotically protected minimal medium, CAA and ara have been included so that a comparison can be made between the β-galactosidase levels in a strain with either low intracellular DAP levels or low intracellular DAP and lysine levels. The results show quite clearly that when extracellular lysine and DAP are removed, there is a much greater level of expression from P_{dapA}. This can be explained by the fact that the intracellular pool of DAP is required for both the synthesis of lysine and incorporation into peptidoglycan, therefore intracellular DAP levels fall more quickly than when it is required for peptidoglycan synthesis alone and consequently P_{dapA} is more rapidly affected.

To show that lysine itself does not affect expression from P_{dapA} a version of MGM100λP_{dapA}-*lacZ* was constructed which had a mutation in *lysA*. A P1 lysate was made using the strain ET505 (*lysA*::*Tn*10) as a donor. The P1 lysate was then used to transduce MGM100λP_{dapA}-*lacZ*. Transductants were selected on appropriate plates containing Tet.

The strain MGM100*lysA*λP_{dapA}-*lacZ* is $dapA^+$ but $lysA^-$. If the strain is grown in medium without lysine it will be able to synthesise DAP but not lysine (and as the internal reserve of lysine is depleted, intracellular lysine levels will decrease). If lysine levels are affecting expression from P_{dapA} then the strain should produce increasing levels of LacZ when grown with no extracellular lysine.

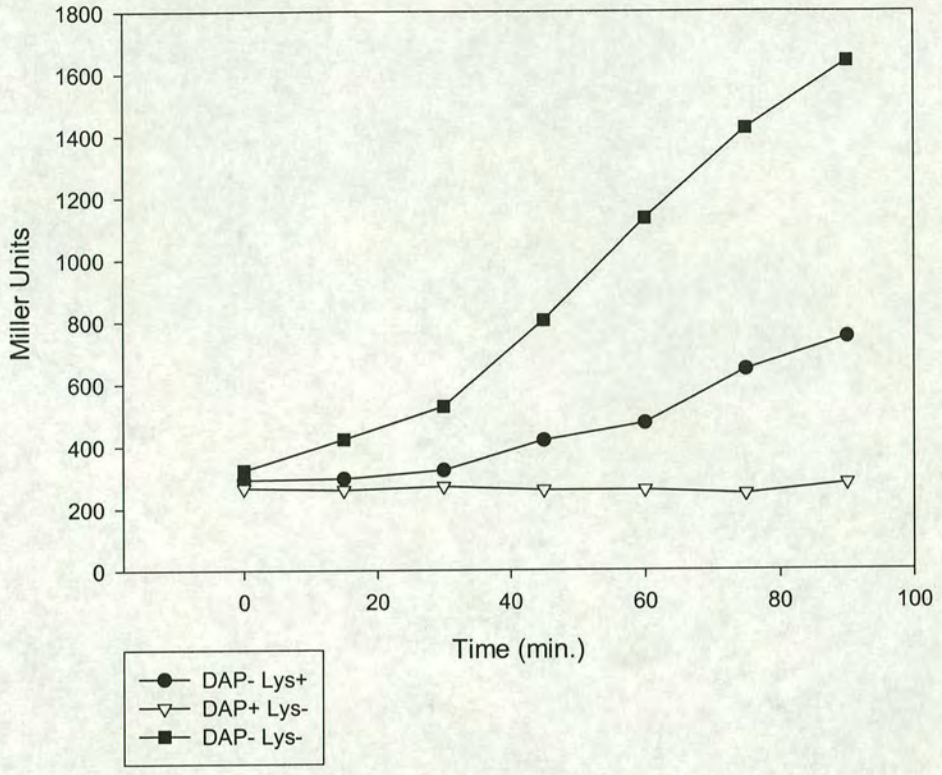


Fig. 5.4.1 β-galactosidase activity in MGM100Δ*dapA*λP_{dapA}-*lacZ* growing in osmotically protected minimal medium, CAA and ara (DAP-, Lys+) or osmotically protected minimal medium, SCAA –Lys, ara with (DAP+, Lys-) or without (DAP-, Lys-) DAP. The strain was grown in minimal medium, CAA, ara and DAP and then centrifuged, the supernatant removed and the cells resuspended in the above media (this was time=0). When intracellular levels of both lysine and DAP are decreased there is a large increase in expression from P_{dapA}.

MGM100*lysA*λP_{dapA}-*lacZ* was grown overnight in minimal medium, CAA, ara, Kan and Tet at 37°C with constant agitation. The following day the culture was diluted 1:100 into 10 ml of fresh medium and grown at 37°C with constant agitation until the OD_{600} was ~ 0.2. 1 ml of culture was removed, centrifuged and the supernatant was removed. The cell pellet was then resuspended in either 10 ml of minimal medium, SCAA –Lys and ara or minimal medium, CAA and ara.

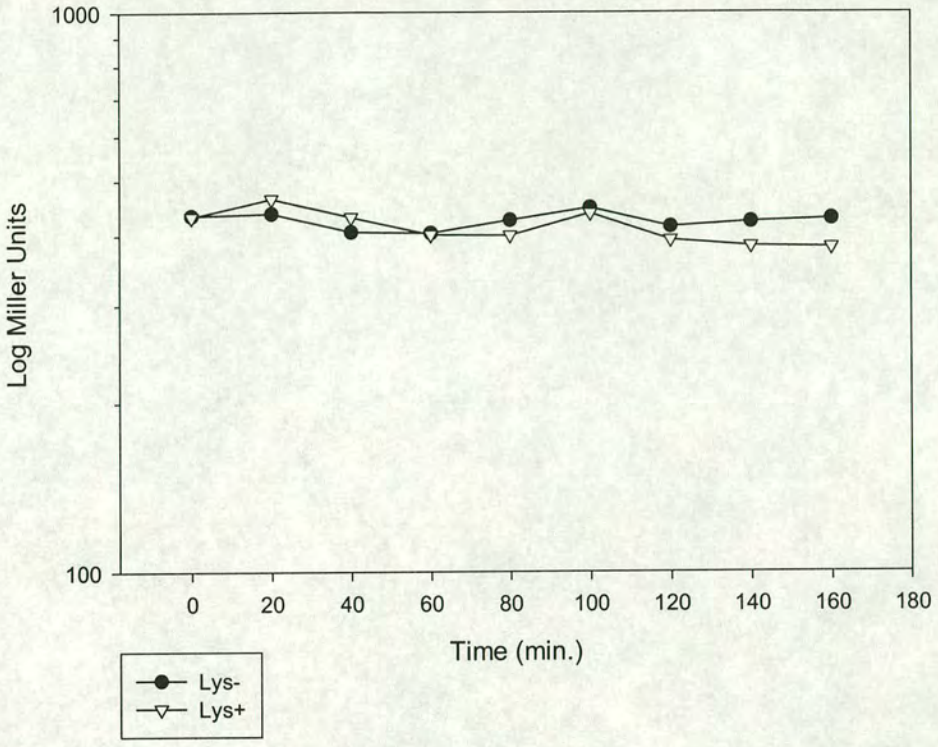


Fig. 5.4.2 β-galactosidase activity in MGM100*lysA*λP_{dapA}-*lacZ* growing in osmotically protected minimal medium, ara and either CAA (Lys+) or SCAA –Lys (Lys-). The strain was grown in minimal medium, CAA and ara and then centrifuged, the supernatant removed and the cells resuspended in the above media (this was time=0). When intracellular levels of lysine are depleted there is no increased expression from P_{dapA}, confirming that lysine does not affect P_{dapA} expression.

The results (Fig. 5.4.2) show that when lysine is removed there is no increased expression from P_{dapA}. This result confirms prior observations that intracellular levels of lysine do not regulate expression from P_{dapA}.

5.5 Growth rate does not affect P_{dapA}

During the course of the previous experiments, it was observed that growth rate during the different growth conditions varied greatly. When MGM100Δ*dapA*λP_{dapA}-*lacZ* was grown in the presence of DAP it had a similar growth rate as MGM100 λP_{dapA}-*lacZ* and it also displayed a similar level of expression from P_{dapA}. But when MGM100Δ*dapA*λP_{dapA}-*lacZ* was grown in the absence of DAP, it had a much slower growth rate (eventually the cells formed spheroplasts and stopped growing) and a higher level of expression from P_{dapA} (this is summarised in Table 5.5.1). Since MGM100 depleted of GroE slows down just before lysis (see chapter 3.1), it was necessary to determine if growth rate affected expression from P_{dapA}.

A P1 lysate was made using the strain CAG18442 (*thrA*::*Tn*10) as a donor. The P1 lysate was then used to transduce MGM100λP_{dapA}-*lacZ*. Transductants (MGM100*thrA*λP_{dapA}-*lacZ*) were selected on appropriate plates containing Tet.

When MGM100*thrA*λP_{dapA}-*lacZ* is grown in minimal medium SCAA –Thr (see chapter 6) it hardly grows at all (any growth is probably due to intracellular stores of threonine), without affecting intracellular levels of DAP. If growth rate, and not DAP, were actually regulating expression from P_{dapA} then there should be an increase in β-galactosidase levels under these conditions.

MGM100*thrA*λP_{dapA}-*lacZ* was grown under the conditions described in section 5.4 (except using SCAA –Thr not SCAA –Lys). The growth curves of MGM100Δ*thrA*λP_{dapA}-*lacZ* in the presence and absence of threonine can be seen in figure 5.5.1 and the results of the β-galactosidase assay can be seen in figure 5.5.2.

Growth Conditions	Growth Rate	Expression from P_{dapA}
$DapA^-$, DAP+	+	+
$DapA^+$, DAP+	+	+
$DapA^+$, DAP-	+	+
$DapA^-$, DAP-	-	++
$DapA^-$, DAP-, Lys-	--	+++

Table 5.5.1 Summary of growth rates and expression from P_{dapA} under the various conditions used. A + sign represents a 'normal' level of growth or expression, multiple + signs represent increased levels and – signs represent decreased levels.

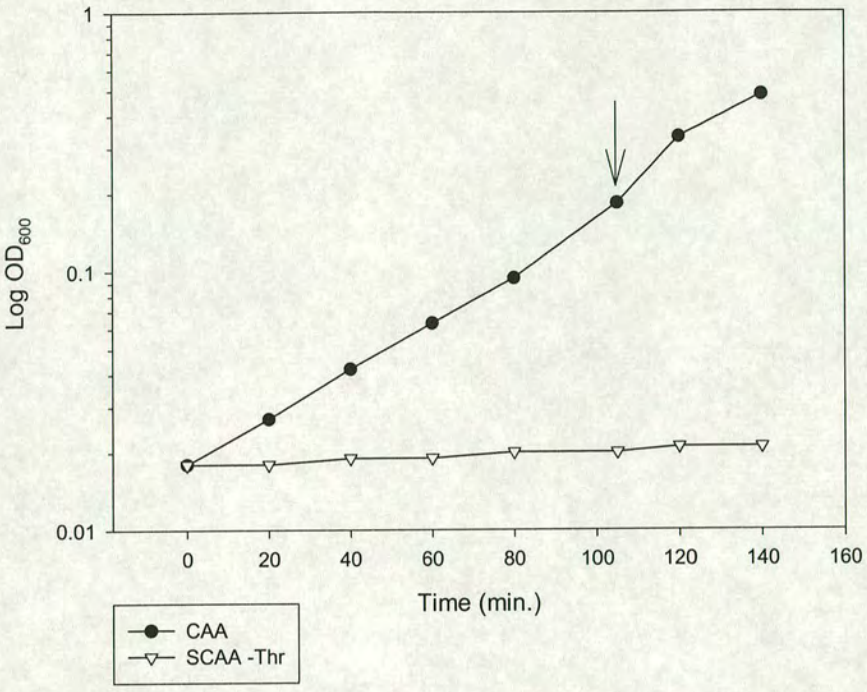


Fig. 5.5.1 Growth rates of MGM100*thrA*λP_{dapA}-*lacZ* growing in osmotically protected minimal medium, ara and either CAA or SCAA –Thr. The strain was grown in minimal medium, CAA and ara and then centrifuged, the supernatant removed and the cells resuspended in the above media (this was time=0). For clarity, when the cells have been diluted 1:10 (delineated by the arrow) the OD has been multiplied by 10 and plotted on the graph.

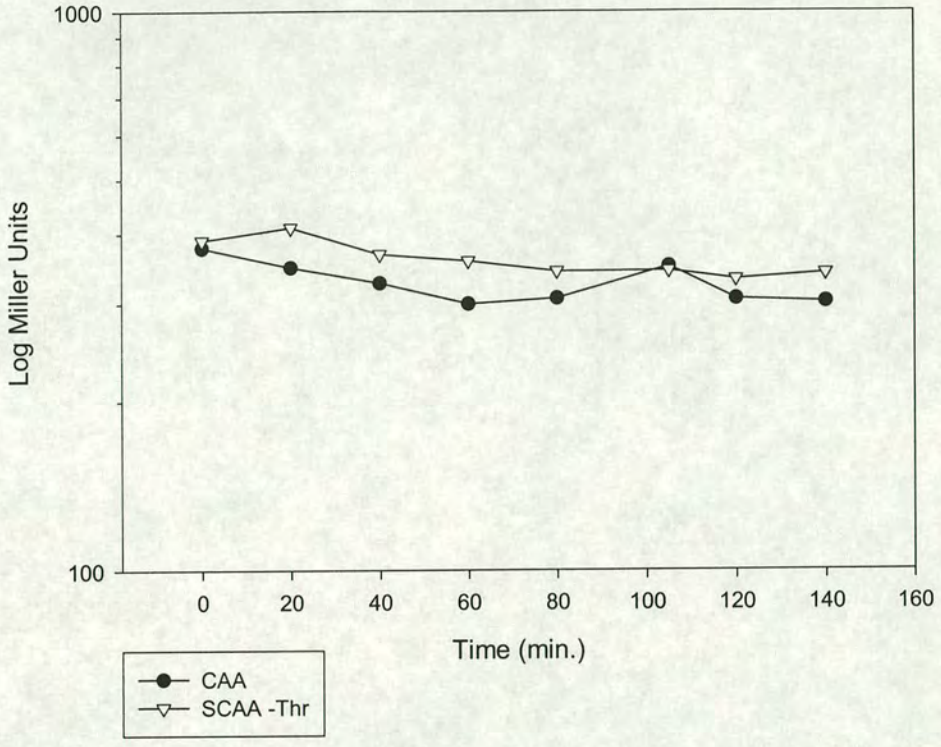


Fig. 5.5.2 β-galactosidase activity in MGM100*thrA*λP_{dapA}-*lacZ* growing in osmotically protected minimal medium, ara and either CAA or SCAA –Thr. The strain was grown in minimal medium, CAA and ara and then centrifuged, the supernatant removed and the cells resuspended in the above media (this was time=0). When growth rate decreases (SCAA –Thr) there is no increased expression from P_{dapA}, confirming that growth rate does not affect P_{dapA} expression.

These results clearly show that even when growth rate is dramatically decreased, there is no increase in the level of expression from P_{dapA}, thereby showing that P_{dapA} is not regulated by growth rate, and is indeed regulated by intracellular levels of DAP.

5.6 Construction of MGM100Δ*dapA*λRSPD72

Since it had now been determined that DAP regulates expression from P_{dapA} it was necessary to determine the method of regulation. The construct MGM100λP_{dapA}-*lacZ* was constructed using ~ 1 kb of DNA upstream of *dapA* ATG (this DNA included the entire coding sequence for the upstream gene *gcvR*). It was hypothesised that P_{dapA} was being regulated by a transcription factor (either an activator or repressor) which was itself affected either directly or indirectly by the intracellular levels of DAP. To determine the binding site for this hypothetical transcription factor it was necessary to construct a new λP_{dapA}-*lacZ* reporter in MGM100 using much less upstream DNA. By steadily decreasing the amount of upstream DNA, and then observing the point at which expression of *lacZ* is unaffected by DAP levels, a putative binding site could be found.

It was decided that the first construct should contain the least amount of upstream DNA that still included the ribosome-binding site and polymerase-binding sites. Two primers were designed, AP1 and AP3, which would produce a 79 bp piece of DNA upstream of the *dapA* ATG. These primers would also introduce a 5′ *Eco*RI and a 3′ *Bam*HI site into the DNA fragment.

AP1:

```
5′  GGGGATCCTCTGTGCAAACA  3′
      BamHI
```

AP3:

```
5′  ATCAGAAGAATTCTGTCTGC  3′
          EcoRI
```

When these primers were used in a PCR using chromosomal DNA as a template, the final product was an 81 bp fragment of DNA which had the following sequence:

```
5′ATCAGAAGAATTCTGTCTGCTTGCTTTTAATGCCATACCAAACGTACCATTGAGA
         EcoRI          -35                      -10

CACTTGTTTGCACAGAGGATCCCC 3′
 mRNA          RBS BamHI
 start site
```

The PCR was performed using an extension time of 30 seconds with MG1655 chromosomal DNA as the template. The PCR product (79 bp) was run on a 1.5% agarose gel, extracted using the β-agarase method and sub-cloned into pGEM-T. The construct was transformed into TOP-10 competent cells and plated on LB + Amp + X-gal, and screened for blue or white colonies. In theory, since the cloning into pGEM-T disrupts the *lacZ* gene, constructs containing the 79 bp fragment should not be able to hydrolyse X-gal and should produce white colonies. In this case it was found that the small insert did not disrupt the *lacZ* gene, and correct constructs resulted in blue colonies.

Three blue colonies were obtained, the plasmid DNA was extracted and subjected to restriction analysis to determine if the construct was correct. One plasmid bearing the correct construct was found (the others lacked inserts). The fragment was released from the plasmid by digesting with the restriction enzymes *Eco*RI and *Bam*HI. The fragment was then ligated to pRS551, which had also been cut with *Eco*RI and *Bam*HI. This produced the plasmid pRS551-PD72 (Fig. 5.6.1). This plasmid now has *lacZ* expression under the control of 70 bp of *dapA* promoter DNA.

pRS551-PD72 was transformed into NM621λRS45. The λRS45 contains the genes *lacA*, *lacY*, *lacZ* and the 5′ terminal half of *bla* (Simons et al, 1987). This allows homologous recombination to occur with the plasmid pRS551 between *bla* and *lacA*. If a double crossover occurs between *bla* and *lacZ-lacA* then the λ will now carry both the Kan^R cassette and *lacZ* under the control of 70 bp of *dapA* promoter DNA.

The strain NM621λRS45 pRS551-PD72 was grown overnight in LB + Kan (to allow recombination to occur) at 37°C with constant agitation. The following day a λ lysate was made (by UV induction) using the protocol described in materials and methods.

The λ lysate was then used to lysogenise MG1655. Lysogens were selected on LB plates containing Kan, and were later screened for Amp^S. Any colonies that were Kan^R and Amp^S were then transduced with a P1 lysate made from MGM100. Since the P_{BAD}*groESL* in MGM100 is linked to a Kan^R cassette it could not be selected with antibiotic resistance; therefore transductants were selected for their ability to grow on LB containing arabinose but not glucose.

The resultant strain (MGM100λRSPD72) was then transduced with a P1 lysate made from MG1655Δ*dapA*. Transductants were selected on plates containing Kan, Cmp and ara.

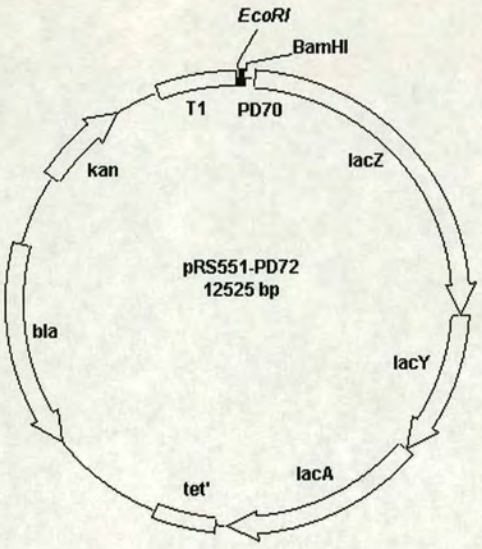


Fig. 5.6.1 The plasmid pRS551-PD72. A 70 bp fragment of DNA upstream of the *dapA* initiation codon (PD70) was cloned into the *Eco*RI and *Bam*HI sites of the plasmid pRS551.

The final strain (MGM100Δ*dapA*λRSPD72) was tested to see if *lacZ* expression was affected by intracellular levels of DAP. The strain was grown as described in section 5.3. The results (Fig. 5.6.2) show that when intracellular levels of DAP decrease, expression of *lacZ* from P_{dapA} increases. This then indicates that the putative transcription factor-binding site is within this 70 bp region of *dapA* promoter DNA (P_{dapA70}).

A computer analysis was run on the 70 bp DNA sequence to determine if any common transcription factor binding sequences were present. None were found. Nor were any inverted repeats found (which might form hairpin loops).

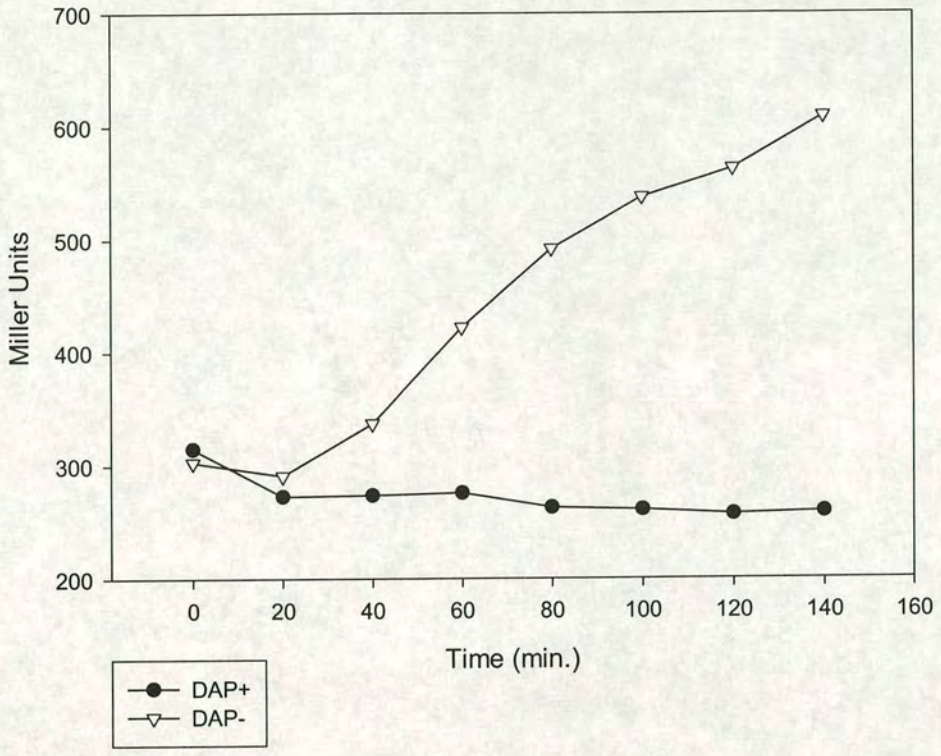


Fig. 5.6.2 β-galactosidase activity in MGM100Δ*dapA*λRSPD72 growing in osmotically protected minimal medium, CAA, ara with (DAP+) or without (DAP-) DAP. The strain was grown in minimal medium, CAA, ara and DAP and then centrifuged, the supernatant removed and the cells resuspended in osmotically protected minimal medium + ara and DAP or just ara (this was time=0). As the intracellular levels of DAP fall in the strain grown without DAP, expression from P_{dapA70} increases.

5.7 Attempts to characterise hypothetical transcription factor

It was important to discover if the hypothetical transcription factor regulating P_{dapA} was a repressor or inhibitor. It was known that when intracellular levels of DAP decreased, expression from P_{dapA} increased, but when excess DAP was present expression did not decrease past a basal level; a $dapA^+$ strain grown with or without extracellular DAP has the same level of *lacZ* expression (see Table 5.2.1).

The strain MGM100Δ*dapA*λRSPD72 was grown as described in section 5.3, and after 60 minutes, DAP was added to the DAP- medium. The results (Fig 5.7.1) show that as intracellular levels of DAP begin to increase again, expression from P_{dapA70} decreases. This suggests that expression from P_{dapA70} increases when intracellular levels of DAP decrease, and decreases as intracellular levels of DAP increase (but only to a basal level, and then no further regardless of the amount of DAP).

To try and determine if the hypothetical transcription factor were an activator or inhibitor, the strain MGM100Δ*dapA*λRSPD72 was transformed with the plasmid TOP-10 PD79 (the plasmid TOP-10 with the 79 bp PCR product formed using primers AP1 and AP3). It was hypothesised that this plasmid would titrate any transcription factor present. If, during DAP depletion, the transcription factor bound to P_{dapA70} and increased transcription of *lacZ* (an activator) then it would also bind to the plasmid TOP-10 PD79, and hence *lacZ* transcription levels would increase less during DAP depletion. If, on the other hand, the transcription factor acted like a repressor, binding to P_{dapA70} in the presence of DAP and prevent transcription of *lacZ* (much like LacI), then the plasmid TOP-10 PD79 should also bind this protein and hence *lacZ* transcription levels would increase more during DAP depletion (also basal transcription in the presence of DAP would be higher).

The strain MGM100Δ*dapA*λRSPD72 transformed with the plasmid TOP-10 PD79 or just TOP-10 was grown overnight in minimal medium, CAA, ara and DAP at 37°C with constant agitation. The following day the culture was diluted 1:100 into fresh medium and allowed to grow at 37°C with constant agitation until OD_{600} was ~ 0.2. Samples were then taken every 10 minutes for 30 minutes and the β-galactosidase activity determined. The results of this experiment are shown in Table 5.7.1. The results indicate that the basal level of *lacZ* expression from P_{dapA70} is unaffected by the presence of extra P_{dapA70} DNA. This would indicate that the hypothetical transcription factor does not act like an inhibitor (if it did, there should have been an increase in basal *lacZ* expression from P_{dapA70} when extra copies of P_{dapA70} DNA were present).

Next the aforementioned strains were grown as described in section 5.3 and the β-galactosidase activity measured in the presence and absence of extracellular DAP. The results (Fig. 5.7.2) show that in the strain containing multiple copies of P_{dapA70} DNA (TOP-10 PD72), when intracellular levels of DAP decrease there is less expression of *lacZ* from P_{dapA70} than in the control strain (TOP-10). This observation can be explained if the hypothetical transcription factor is acting like an activator.

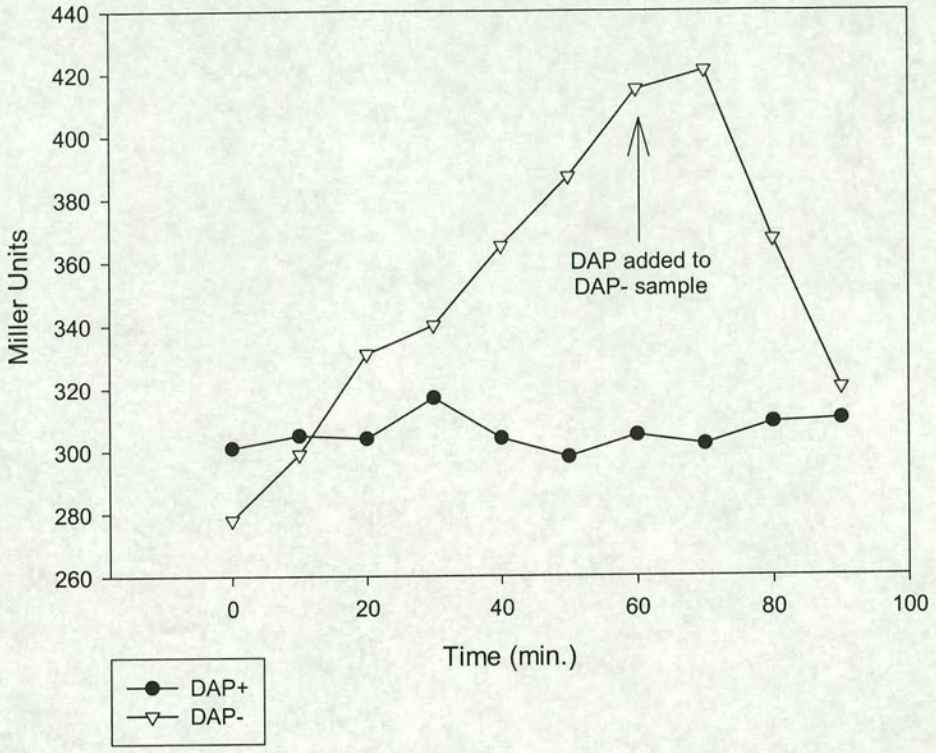


Fig. 5.7.1 β-galactosidase activity in MGM100Δ*dapA*λRSPD72 growing in osmotically protected minimal medium, CAA, ara with (DAP+) or without (DAP-) DAP. The strain was grown in minimal medium, CAA, ara and DAP and then centrifuged, the supernatant removed and the cells resuspended in osmotically protected minimal medium + ara and DAP or just ara (this was time=0). At 60 minutes, DAP was added to the DAP- sample. As the intracellular levels of DAP increase again in the strain grown without DAP, expression from P_{dapA70} decreases.

Strain	β-gal #1	β-gal #2	β-gal #3	β-gal #4	Ave β-gal
MGM100Δ*dapA*λRSPD72 TOP-10 PD72	289.7	300.7	298	303.3	297.9
MGM100Δ*dapA*λRSPD72 TOP-10	287.6	286.6	304.3	297.7	294.08

Table 5.7.1 The strain MGM100Δ*dapA*λRSPD72 containing either the plasmid TOP-10 PD72 or TOP-10 was grown in minimal medium, CAA, ara and DAP. Samples were taken every 10 minutes for 30 minutes and the β-galactosidase activity determined; this was repeated for four different cultures. The average of these three readings are recorded as β-gal #1-4. The final average of cultures 1-4 (a final average of 12 readings, over 4 cultures) is recorded as Ave β-gal. The results indicate that basal expression of *lacZ* from P_{dapA70} is unaffected by the presence of extra copies of P_{dapA70} DNA.

When the cells have a high level of intracellular DAP, the hypothetical transcription factor may not bind to P_{dapA70} and therefore there is only a basal level of expression of *lacZ*. Under these circumstances, adding excess DAP would not cause a decrease in expression from P_{dapA70}. When the intracellular levels of DAP decrease, the transcription factor might then bind to P_{dapA70} and cause an increased expression from the *dapA* promoter.

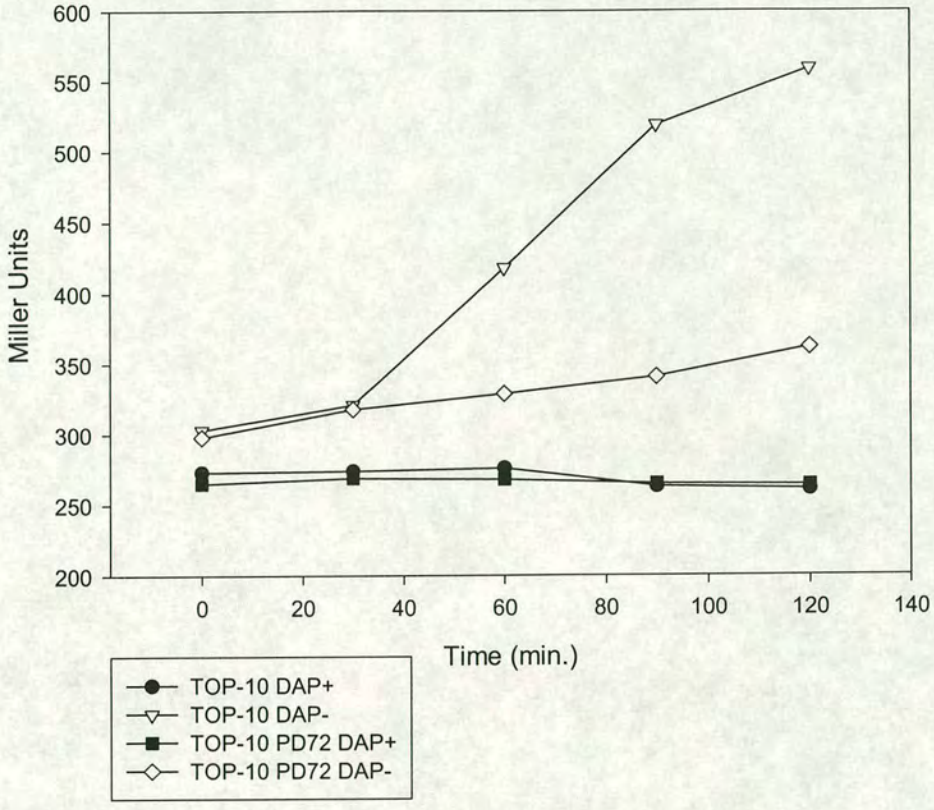


Fig. 5.7.2 β-galactosidase activity in MGM100Δ*dapA*λRSPD72 containing either the plasmid TOP-10 or TOP-10 PD72, growing in osmotically protected minimal medium, CAA, ara with (DAP+) or without (DAP-) DAP. The strains were grown in minimal medium, CAA, ara and DAP and then centrifuged, the supernatant removed and the cells resuspended in osmotically protected minimal medium + ara and DAP or just ara (this was time=0). As the intracellular levels of DAP fall in the strain grown without DAP, but containing multiple copies of P_{dapA70} DNA (TOP-10 PD72 DAP-), expression from P_{dapA70} does not increase at the same rate as it does in the strain containing only 1 copy of P_{dapA70} (TOP-10 DAP-). This indicates that the transcription factor acts like an activator.

5.8 Gel retardation assays on P_{dapA70}

In an attempt to identify the hypothetical transcription factor, a gel retardation assay was performed using the P_{dapA70} fragment and lysates from MGM100Δ*dapA* grown in the presence or absence of DAP.

Gel retardation is a powerful approach for both qualitative and quantitative analysis of protein-nucleic acid interactions. The method is based on the observation that the electrophoretic mobility of a nucleic acid through a polyacrylamide gel can be altered when a protein is bound to it (Garner et al, 1981; Fried et al, 1981). The gel method is ideal for purifying new proteins that bind to a defined target DNA. Two approaches can be taken: the gel method can be used to assay classical biochemical purification steps, or the protein can be purified directly from the complex in the gel.

In crude cell extracts, detection of binding activity toward a specific target DNA is complicated by the presence of other binding proteins. It is possible to detect a specific binding protein against this background of non-specific binding activities because the specific and non-specific proteins are competing for binding to the target DNA, but the specific complex has a higher affinity under most conditions. The addition of non-specific DNA (poly dIdC) and in this instance, non-specific promoter DNA (pUC18) should cause a decrease in non-specific binding to the DNA fragment of interest.

The P_{dapA70} DNA fragment (generated by PCR using primers AP1 and AP3) was labelled with [γ-^{32}P] ATP using the T4 polynucleotide kinase reaction as described in materials and methods.

The strain MGM100Δ*dapA* was grown in osmotically protected minimal medium, CAA, ara and either the presence or absence of DAP. 100 ml of the culture was grown at 37°C with constant agitation until the OD_{600} was ~ 0.5. The culture was then centrifuged for 5 minutes at 1500 x g and the supernatant was discarded. The cell pellet was then resuspended in 1 ml of 0.1M Tris-Cl (pH 7.4) and the cells lysed by sonication. The lysates were then subjected to a further centrifugation of 14000 x g for 5 minutes, and the supernatant (soluble cell lysate) removed for future use.

For the band shift experiment, 10 pmol of ^{32}P-P_{dapA70} was incubated with 400 ng of pUC18, 5 μg of poly dIdC, 1x band shift buffer (20mM Tris-Cl (pH 7.9), 1mM EDTA, 50mM, NaCl) and varying amounts of either DAP+ or DAP- lysates. The incubations were carried out at 37°C for 15 minutes. The mixtures then had glycerol added at a final concentration of 5% (v:v) and were loaded onto a 6% (w:v) acrylamide DNA sequencing gel, without urea (see materials and methods).

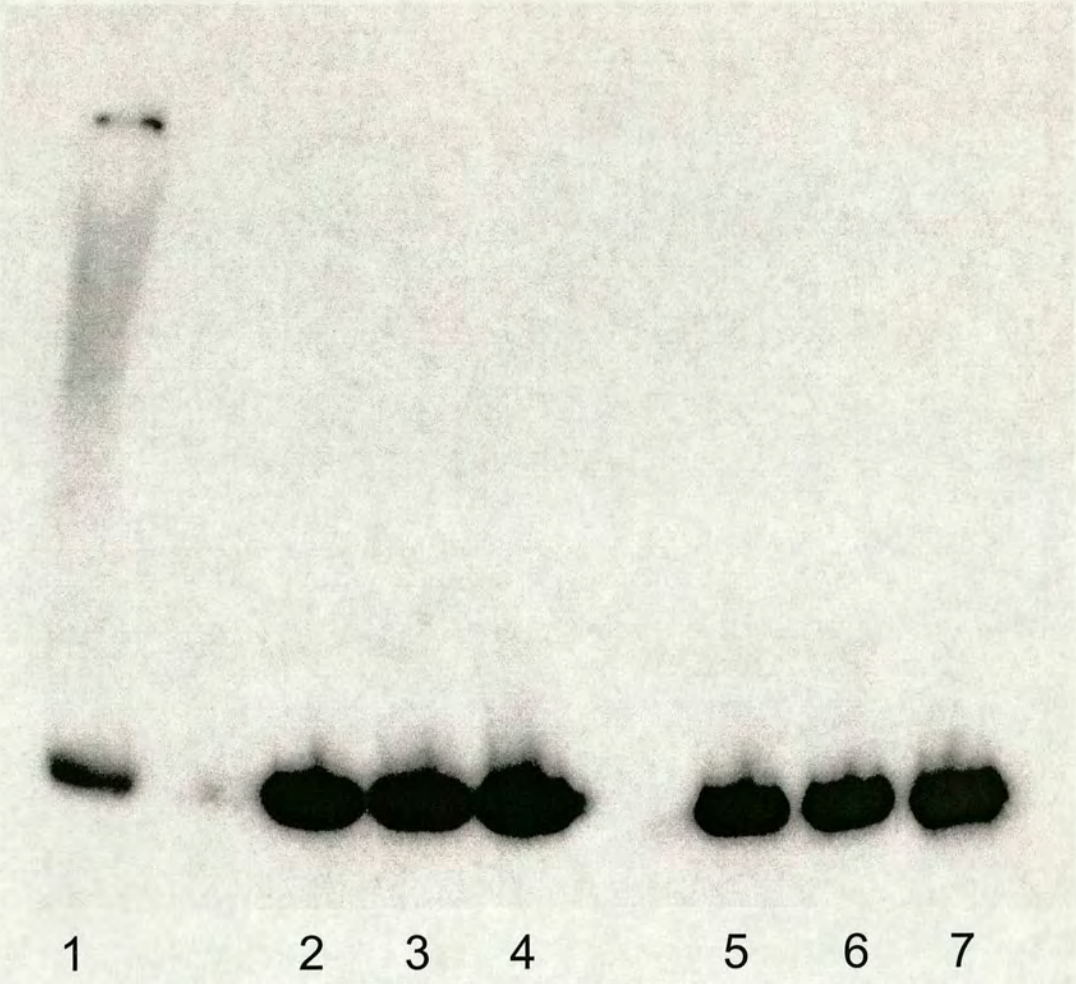


Fig. 5.8.1 Gel retardation assay on P_{dapA70}. P_{dapA70} DNA was incubated with 0.1μl, 1μl or 10μl of either DAP+ lysate (lanes 2-4 respectively) or DAP- lysate (lanes 5-7 respectively) under the conditions described in the text. Lane 1 is a control lane which has neither lysate added. There is no band shift in the presence of either lysate at any concentration.

The results (Fig. 5.8.1) show that there was no band shift using either a DAP+ or DAP- lysate at varying concentrations. Though this experiment was repeated many times, using widely-ranging concentrations of both ^{32}P-P_{dapA70} DNA and lysates, no band shift was ever observed. There are a number of possible explanations as to why this might have been: the protein may have been inactive, too little protein may have been present or the gel pH may have been too high.

5.9 Yeast one hybrid system

In a further attempt to isolate the hypothetical transcription factor responsible for the regulation of the *dapA* promoter, a yeast one-hybrid system was used (Grow 'n' Glow, MoBiTec GmbH).

The yeast one-hybrid system (Wang et al, 1993), a modified version of the two-hybrid system, rapidly identifies DNA binding proteins from either cDNA libraries or known gene sequences. The one–hybrid system exploits the bimodular structure of eukaryotic transcriptional activators. A DNA binding element of interest is cloned into the promoter region controlling expression of the green fluorescent protein (GFP) reporter in the GFP reporter vector. The particular GFP variant expressed by the GFP reporter plasmid, GFPuv, has the same excitation and emission maxima as wild-type GFP, but is 18 times brighter than the wild-type variant (Crameri et al, 1996). A gene encoding a potential DNA binding protein is cloned in the prey plasmid so that it is expressed as a fusion to a B42 activator peptide. Interaction of the DNA binding protein with the DNA element brings the B42 activator peptide into a position that initiates transcription of the GFP reporter gene. Thus, colonies with DNA binding proteins that interact with the DNA element of interest glow bright green, which is easily seen by eye under UV illumination (300 nm) (Fig. 5.9.1).

Expression of fusion proteins by the prey plasmid is controlled by the GAL1 upstream activation sequence. In yeast with an intact galactose regulatory system, the GAL1 activation sequence is induced by galactose and repressed by glucose. This regulation provides a mean to prevent expression of library/activation domain fusions until the actual screening, which avoids any potentially toxic library proteins affecting growth of the yeast.

Yeast is a suitable organism for this work as the lysine biosynthetic pathway in yeast is very different from the lysine biosynthetic pathway in bacteria. The commonly known pathway for the biosynthesis of lysine in bacteria is initiated by the condensation of pyruvate and aspartic semialdehyde to form dihydropicolininc acid and involves diaminopimelic acid as a key intermediate (Davis, 1952). The second pathway for the biosynthesis of lysine present in fungi and blue-green algae (cyanobacteria) begins with a condensation of acetate and α-ketogluterate to form homocitric acid and involves α-aminoadipic acid as a key intermediate (Gilvarg et al, 1951). Since yeast does not utilise DAP there is little chance that it might produce any DAP regulated transcription factors that would bind to the introduced *E. coli* DNA 'bait' sequence.

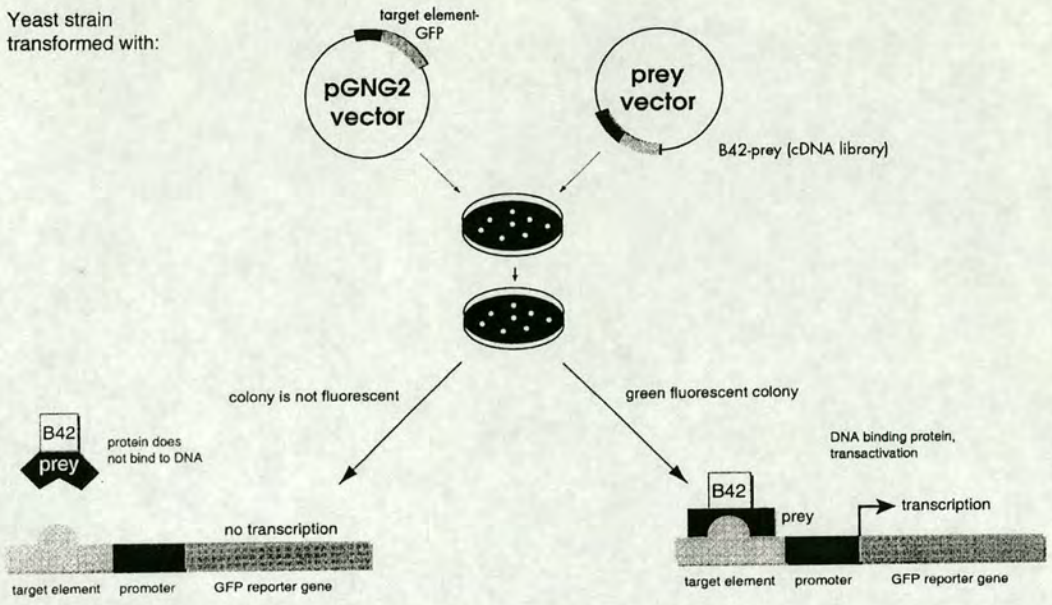


Fig. 5.9.1 Overview of the yeast one-hybrid system. A fusion protein (DNA binding protein-B42 activator peptide) that binds to the DNA region of interest (target element) will activate expression of GFPuv, resulting in bright green colonies when expose to light of a wavelength of 300 nm.

5.9.1 Construction of a DNA library

The plasmid pJG4-5 was cut with the restriction enzyme *Eco*RI and treated with calf intestinal phosphatase (CIP) to prevent religation. 20 μg of chromosomal DNA from MG1655 was cut with 20 units of the restriction enzyme *Apo*I for 10 minutes at 37°C (these conditions gave the greatest number of fragments with sizes between 1.5 and 4 kb). The partially digested chromosomal DNA was then ligated to pJG4-5 using the protocol described in materials and methods. The ligated plasmid/chromosomal DNA was then transformed into TOP-10 and selected on plates containing Amp. All Amp^R colonies were pooled together and their plasmid DNA extracted using the protocol outlined in materials and methods.

5.9.2 Construction of pGNG-72

The plasmid pGNG-2 was extracted from the strain INV110, cut with the restriction enzyme *Nru*I and treated with CIP to prevent religation. The DNA bait

region (P_{dapA70}) was generated by PCR using primers AP1 and AP3, but using Vent DNA polymerase, to give a blunt ended product. The PCR product was run on an agarose gel, the correct size band extracted using the β-agarase method, and ligated to *Nru*I cut pGNG-2. Ligated plasmid/DNA was transformed into TOP-10 and colonies selected on LB plates containing Amp. A number of colonies were picked, the plasmid DNA extracted using the protocol described in materials and methods, and the plasmid DNA sequenced to determine if any had the bait sequence ligated to the plasmid. A correct plasmid was found and called pGNG-72.

5.9.3 Screening of one-hybrid system

The plasmid pGNG-72 was transformed into the yeast strain EGY48 (using the Glow 'n' Grow high efficiency yeast transformation kit, MoBiTec) and transformants were selected for their ability to grow on DOBA –Ura (Glucose) (EGY48 has a mutation in the URA3 gene while the plasmid pGNG-72 carries an intact copy of the URA3 gene, and can therefore complement the mutation, allowing growth in the absence of uracil).

EGY48 pGNG-72 was then transformed with a large amount of the plasmid pJG4-5 containing the chromosomal DNA library. This time transformants were selected for their ability to grow on DOBA –Ura –Trp (Galactose) (EGY48 also has a mutation in the TRP1 gene, while pJG4-5 carries an intact version of the TRP1 gene). Since the plates contained galactose, expression of the fusion library was now active, and colonies were screened for any positives (glowed bright green when exposed to UV light of 300 nm).

Although this experiment was repeated many times, as was the construction of the chromosomal DNA library, no positives were ever found.

5.10 Summary and discussion

In this chapter work has been described that suggest the possibility of a novel transcription factor for *dapA* in *E. coli*.

The original observation that led to this discovery was that during GroE depletion the expression of *dapA* increased. Until this time it had been thought that expression of *dapA* was not regulated (Butour et al, 1974; Bouvier et al, 1984). During the course of the work it was discovered that this regulation was not by the peptide DapA (a Δ*dapA* strain grown in the presence of DAP had the same level of expression from the *dapA* promoter as a *dapA*$^+$ strain) but by one of its downstream products: DAP.

In *E. coli* the lysine biosynthetic pathway is not reversible except at the step catalysed by DapF (the racemisation of LL-diaminopimelate to meso-diaminopimelate), therefore when DAP is added to a Δ*dapA* strain, only the level of those metabolites that are downstream of the DapE catalysed step are affected, i.e. the addition of meso-DAP will only change the levels of LL-DAP, meso-DAP and lysine. Therefore when DAP was added to DAP depleted cells and observed to change the level of expression from the *dapA* promoter, only one of the aforementioned compounds could have been responsible.

When *E. coli* were depleted of their intracellular stores of lysine in a Δ*dapA* strain, expression from P_{dapA} was observed to increase dramatically, though when this experiment was repeated in a *lysA* mutant strain, the effect was not observed at all. This observation could be explained by the fact that in the first instance DAP was being exhausted as it was required for both the synthesis of lysine and peptidoglycan, but in the second instance (the *lysA* mutant) DAP could not be converted into lysine, and therefore DAP levels were unaffected in this experiment. As expression from P_{dapA} only changed when intracellular DAP levels changed, this was good evidence that DAP was actually regulating expression from P_{dapA}.

To make sure that changes in the intracellular levels of DAP were only affecting expression from the *dapA* promoter and did not affect *lacZ,* experiments were conducted in which the chromosomal levels of *lacZ* (under its native promoter) were measured during DAP depletion (MGM100Δ*dapA* was grown under conditions which caused DAP depletion). In these experiments chromosomal *lacZ* expression was unaffected by changes in the intracellular levels of DAP. This showed that DAP levels do not affect *lacZ* expression unless it is under the control of the *dapA* promoter.

In later work the region in which the transcription factor could bind was narrowed down greatly from ~ 1 kb of DNA to ~ 70 bp of DNA. The P_{dapA70} reporter construct acted in the same way as the previous reporter construct. Unfortunately the binding region could not be narrowed down any further using the truncation method as to do so would incur the loss of the RNA polymerase binding sites. It was hoped that if a transcription factor had been discovered, its exact binding region could have been determined using *in vitro* DNA footprinting.

Two methods were used to try and discover the identity of the transcription factor, band shift and yeast one-hybrid. The band shift proved inconclusive, as the mobility of the 70 bp DNA strand was not affected by the presence of a cell lysate made from *E. coli* either with or without normal intracellular levels of DAP. But, if the transcription factor was present only at very low amounts (as may be expected) then it is possible that a mobility shift may not have been observed. Also, if the conditions were not correct for the binding of the transcription factor to the DNA fragment, again no mobility shift would be observed. There is also the possibility that in the DAP depleted cell lysate, the transcription factor is tightly bound to the chromosomal *dapA* promoter region and none is actually present in the lysate at all! The gel retardation assay might be better employed to verify that purified protein does actually bind to the DNA region of interest.

The yeast one-hybrid system is an elegant approach to discovering the identity of proteins that interact with a DNA region of interest. Unfortunately in this case a pre-made *E. coli* chromosomal library in the plasmid pJG4-5 was not available. This would have made the entire process much easier, as the library construction was probably the limiting step in this experiment. If the fusion library does not give representative coverage of the chromosome then it may be unlikely that any DNA binding proteins of interest would be found.

Another approach to purifying the transcription factor could be to use DAP to purify the protein. If DAP were immobilised to an appropriate resin (sephadex for example), then cell lysates could be passed over this resin to see what binds to the DAP. Of course the problems with this approach are numerous, how tight is the interaction between DAP and the protein? Under what conditions do they interact? Under what conditions do they dissociate? And most importantly, does DAP interact directly with the transcription factor, or does it do so via a second protein (i.e. DAP binds to a protein, which in turn modifies the transcription factor, causing it to change expression of the *dapA* promoter)?

There is of course the possibility that *dapA* is not regulated at the level of transcription, but at the level of translation. In the P_{dapA}-*lacZ* reporter constructs, the

first 21 bases of the fusion mRNA is *dapA* mRNA. It is possible that this small region of mRNA may contain some sequence, which allows for the regulation of its translation (though this sequence did not contain any hairpin-loop forming regions, which tend to characterise regions controlling regulation of translation). Probably the easiest way to determine if *dapA* is being regulated at the level of translation is to measure the amount of *dapA-lacZ* fusion mRNA during DAP starvation. If the level of mRNA does not increase, but LacZ activity does increase, then this is good evidence that expression from the *dapA* promoter is being regulated at the translational level.

In conclusion, in this chapter it has been proposed that the expression of *dapA* is regulated in a hitherto unknown manner by the metabolite, DAP. Though this work is still incomplete, the identification of a *dapA* DNA binding protein, or a protein that regulates translation of *dapA* mRNA in response to fluctuating DAP levels, may help in our further understanding of the regulation of genes in *E. coli*.

Chapter VI

Growth of MGM100 in amino acid deficient media

6.1 Introduction

In initial studies of DapA it was thought that the protein had an increased rate of turnover in GroE depleted cells. To determine if this were true, newly synthesised DapA was labelled with ^{35}S-methionine in either GroE^{+} or GroE^{-} cells, an excess of unlabelled methionine was added, and samples were taken at various times to determine how much radiolabelled DapA remained.

Unfortunately, uptake of ^{35}S-methionine in LB is very low (due to the presence of methionine in the medium), and while the cells could be labelled in minimal medium, they exhibited a much slower growth rate in this medium. Therefore it became necessary to formulate a new medium; one in which cells would have a similar growth rate as in LB but which also allowed sufficient uptake of ^{35}S-methionine to allow efficient labelling of newly synthesised proteins.

6.2 Radiolabelling of MGM100

In earlier experiments (see chapter 3) it was necessary to radiolabel newly synthesised proteins using ^{35}S-methionine. Until this time MGM100 had been consistently grown in LB, but ^{35}S-methionine uptake in this medium is very low (due to the presence of methionine in the medium, see figure 6.2.1). Therefore it was necessary to perform radiolabelling of MGM100 in a medium that lacked methionine. One approach was to grow MGM100 in Spizizen's minimal medium (MM); while growth in this medium allowed the highest uptake of ^{35}S-methionine (Fig. 6.2.1) the medium was rejected as the cells had a vastly different generation time to those grown in LB (generation time in LB and ara is ~ 22 minutes, while in MM and ara it is ~ 60 minutes). Another approach was to grow MGM100 in LB and then rapidly filter the cells and resuspend them in MM just before the labelling process. While this method gave better ^{35}S-methionine uptake than did LB (Fig. 6.2.1) the uptake was still much less than when the cells were grown in MM alone. In all of the above experiments the cells were grown in the appropriate medium overnight at 37°C with constant agitation, the following day the cells were diluted 1:100 into fresh medium and grown at 37°C with constant agitation until the OD_{600} was ~ 0.2. The culture was then diluted 1:10 into fresh prewarmed medium and grown for 60 minutes. After 60 minutes ^{35}S-methionine was added at a final activity of 100 μCi ml^{-1}, this was time 0 (for the sample that was shifted from LB into MM,

the cells were vacuum filtered through 0.45 μm Whatman filters and resuspended in prewarmed MM just before the radiolabel was added). 10 ml of culture at OD_{600} 0.2 was removed at various times, the cells were vacuum filtered through 0.45 μm Whatman filters, washed with an ice cold solution of 5% trichloroacetic acid (to remove surface macromolecules) and the filters were dried. The filters were then put into scintillation fluid, and the counts per minute (C.P.M.) were determined on a scintillation counter.

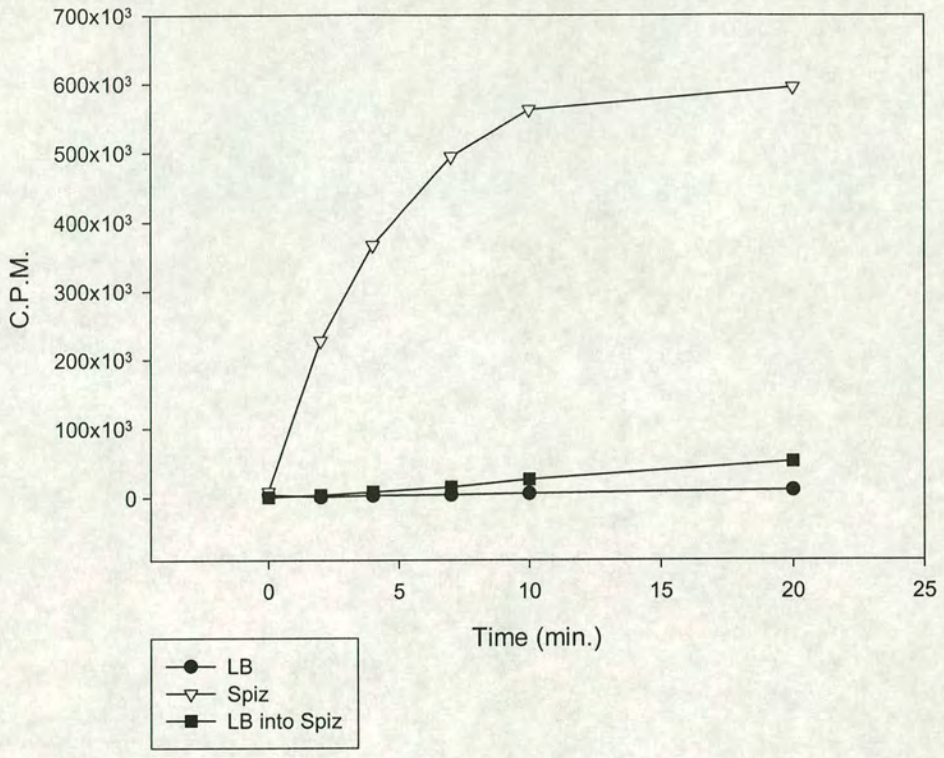


Fig. 6.2.1 Scintillation counts of MGM100 labelled with 100 μCi ml^{-1} ^{35}S-methionine in various media. Cells were grown in LB (LB), Spizizen's minimal medium (Spiz) or LB and then shifted to Spizizen's minimal medium just before adding the radiolabel (Spiz into LB). Growth conditions are described in the text. In both the LB and LB into Spiz samples, uptake of ^{35}S-methionine is very low compared to the Spiz sample. The plateau In the Spiz sample is not caused by a limiting amount of radiolabel (there was a total of $1.4x10^7$ CPM available).

6.3 Composition of SCAA

It was noted that the generation time of MGM100 growing in MM supplemented with casamino acid (CAA) was ~ 30 minutes. This was much closer to the generation time in LB and it was therefore decided to try and create a supplement which had the same composition as CAA but lacked methionine.

Casamino acid is formed by the acid hydrolysis of casein. Before a supplement that had the same composition as CAA but lacked methionine could be created, it was necessary to determine the composition of CAA.

Casein (from Bovine milk) is composed of four different types of casein, α_{s1}, α_{s2}, β and κ, which account for 39.7%, 10.3%, 36.9% and 13.1% of the total casein respectively (Swaisgood, 1992). The amino acid composition of the different forms of casein was determined (Swaisgood, 1992) and can be seen in Table 6.3.1. Using this information the approximate composition of CAA (at a concentration of 100 mg ml^{-1}) could be determined (Table 6.3.2).

Growth curves were plotted to determine if MM supplemented with the synthetic casamino acid (SCAA) solution gave the same generation time as MM supplemented with CAA. MGM100 was grown at 37°C in MM1 [MM supplemented with thymidine, adenine, uracil, tryptophan, vitamin B1 and either CAA (5 mg ml^{-1}) or SCAA (5 mg ml^{-1})]. The OD_{600} was measured at various times and plotted. The results (Fig. 6.3.1) show that MGM100 grown in MM1 SCAA and ara has the same generation time as MGM100 grown in MM1 CAA and ara.

Since MM SCAA appeared to supply the same nutrients as MM CAA, a new version was made from which methionine was omitted (SCAA –Met). When MGM100 was grown in MM1 SCAA –Met and ara it was found to have a generation time of ~ 33 minutes, but it also allowed for a much greater uptake of ^{35}S-methionine (Fig. 6.3.2). This demonstrated that MM1 SCAA –Met was a suitable medium for use in radiolabelling experiments with ^{35}S-methionine.

Casein	α_{s1}			α_{s2}			β			κ		
Amino Acid	No.	%	Conc.	No.	%	Conc.	No.	%	Conc.	No.	%	Conc.
Ala	9	4.7	1.87	8	4.1	0.42	14	9	3.31	5	2.5	0.32
Arg	6	3.1	1.24	6	3.1	0.31	5	3	1.10	4	2	0.26
Asn	8	4.2	1.66	14	7.1	0.73	8	4.8	1.76	5	2.5	0.32
Asp	7	3.7	1.45	4	2	0.21	3	1.8	0.66	4	2	0.26
Cys	0	0	0	2	1	0.10	2	1.2	0.44	0	0	0
Gln	14	7.3	2.90	16	8.2	0.84	14	8.4	3.09	20	9.8	1.28
Glu	25	13	5.19	24	12.2	1.26	12	7.2	2.65	19	9.3	1.22
Gly	9	4.7	1.87	2	1	0.10	2	1.2	0.44	5	2.5	0.32
His	5	2.6	1.03	3	1.5	0.15	3	1.8	0.66	5	2.5	0.32
Ile	11	5.8	2.28	11	5.6	0.57	13	7.8	2.87	10	4.9	0.64
Leu	17	8.9	3.53	13	6.6	0.68	8	4.8	1.76	22	10.8	1.41
Lys	14	7.3	2.90	24	12.2	1.26	9	5.4	1.98	11	5.4	0.71
Met	5	2.6	1.03	4	2	0.21	2	1.2	0.44	6	2.9	0.38
Phe	8	4.2	1.66	6	3.1	0.31	4	2.4	0.88	9	4.4	0.58
Pro	17	8.9	3.53	10	5.1	0.52	20	12	4.41	35	17.2	2.25
Ser	8	4.2	1.66	6	3.1	0.31	12	7.2	2.65	11	5.3	0.71
Thr	5	2.6	1.03	15	7.7	0.78	14	8.4	3.09	9	4.4	0.58
Trp	2	1	0.41	2	1	0.10	1	0.6	0.22	1	0.5	0.06
Tyr	10	5.2	2.07	12	6.1	0.63	9	5.4	1.98	4	2	0.26
Val	11	5.8	2.28	14	7.1	0.73	11	6.6	2.43	19	9.3	1.22
Total	**191**	**100**	**39.7**	**196**	**100**	**10.3**	**167**	**100**	**36.9**	**204**	**100**	**13.1**

Table 6.3.1 The amino acid composition of the four types of casein. The number of residues of each amino acid is shown (No.) as well as the percentage of the casein that they represent (%). To make a 100 mg ml^{-1} solution of SCAA with the same amount of each amino acid as CAA, α_{s1} casein must be added at 39.7 mg ml^{-1}, α_{s2} casein must be added at 10.3 mg ml^{-1}, β casein must be added at 36.9 mg ml^{-1} and κ casein must be added at 13.1 mg ml^{-1}. The columns marked Conc. show how much of each amino acid must be present (in mg ml^{-1}) to make up the appropriate amounts of the individual caseins.

Amino Acid	mg/100 ml	Amino Acid	mg/100 ml
Ala	592.65	Leu	739.71
Arg	292.41	Lys	686.62
Asn	448.73	Met	207.67
Asp	258.49	Phe	343.99
Cys	54.70	Pro	1072.57
Gln	812.85	Ser	533.60
Glu	1032.92	Thr	549.89
Gly	273.88	Trp	80.60
His	218.09	Tyr	495.46
Ile	637.91	Val	667.27

Table 6.3.2 The composition of SCAA. The Conc. values from Table 6.3.1 were totalled and multiplied by 100 for each amino acid. By adding the amino acids above at the noted concentrations a 100 ml solution of SCAA (at a concentration of 100 mg ml^{-1}) will be obtained which will have the same amino acid composition as CAA at a concentration of 100 mg ml^{-1}.

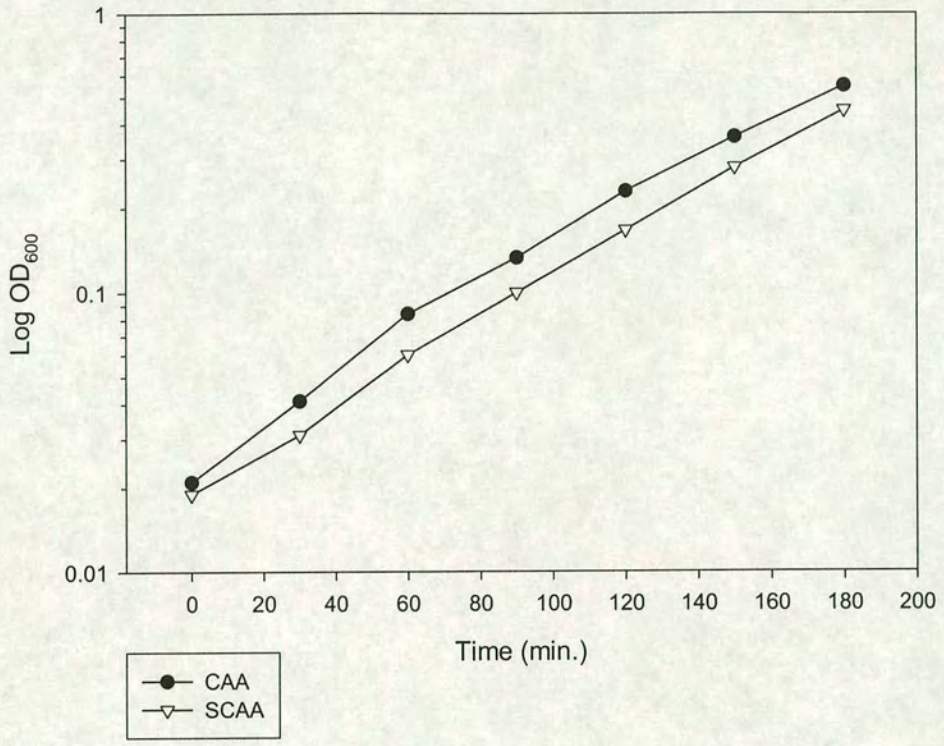


Fig. 6.3.1 Growth curves of MGM100 grown in MM1 supplemented with either CAA (5 mg ml^{-1}) and ara or SCAA (5 mg ml^{-1}) and ara. MGM100 was grown overnight in 5 ml of the appropriate medium at 37°C with constant agitation. The following day the culture was diluted 1:100 into fresh medium and grown at 37°C with constant agitation until the OD_{600} was ~ 0.2. The culture was then diluted 1:10 into fresh, prewarmed medium (this was time = 0) and grown at 37°C with constant agitation. The OD_{600} was measured every 30 minutes. For clarity, where the culture has been diluted 1:10, the OD has been multiplied by 10 and plotted on this graph. Both culture were diluted 1:10 at 120 minutes.

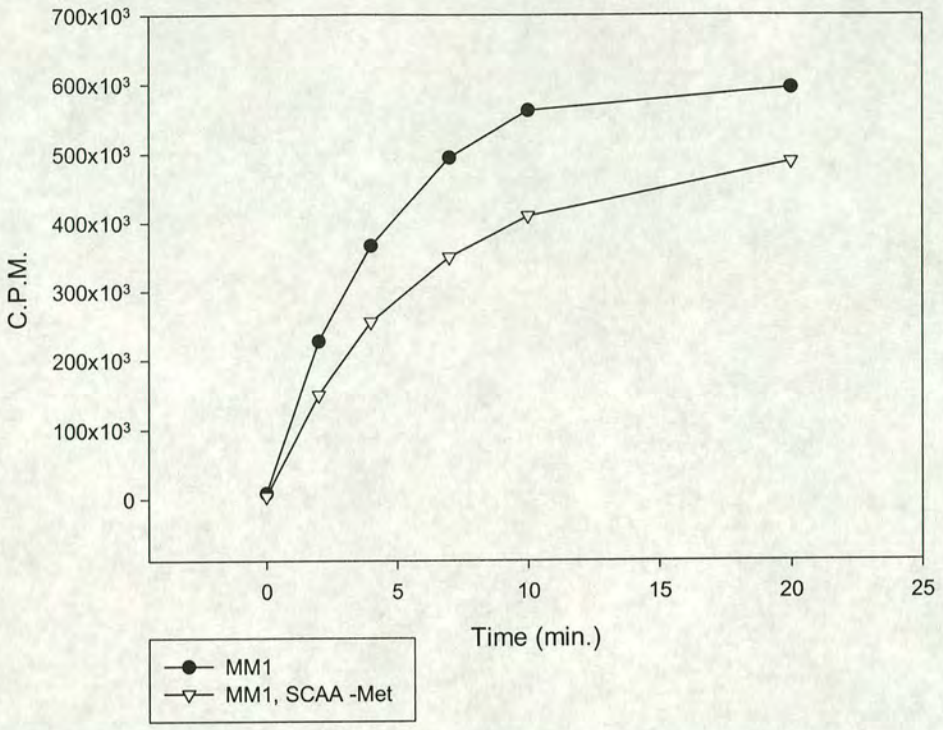


Fig. 6.3.2 Scintillation counts of MGM100 labelled with 100 μCi ml^{-1} ^{35}S-methionine in various media. Cells were grown in MM1 or MM1 supplemented with SCAA –Met (5 mg ml^{-1}) at 37°C with constant agitation. 100 μCi ml^{-1} ^{35}S-methionine was added (this was time = 0) and samples were taken at various times to determine the uptake of the radiolabel. MGM100 grown in MM1 supplemented with SCAA –Met has a similar uptake level as MGM100 grown in un-supplemented medium.

When newly synthesised proteins in $GroE^+$ cells were labelled with ^{35}S-methionine in MM1 SCAA –Met, very good results were obtained (Fig. 6.3.3); but when labelling was attempted in GroE depleted cells, the radiolabelling was much less efficient (Fig. 6.3.3). To explain this observation it was suggested that methionine uptake might be affected by GroE depletion. To visualise radiolabelled proteins in GroE depleted cells, a 2D gel was used to separate proteins; the proteins were then transferred to a PVDF membrane using the Western blotting technique (described in materials and methods) and the membrane was covered with a thin layer of EA wax (EA Biotech Ltd.). EA wax contains a scintillant, which amplifies

the signal from ^{35}S. When the membrane is exposed to X-ray film, a much clearer image can be obtained than if the wax were not used (using EA wax it was possible to obtain an image in 3 days on X-ray film that could not be seen after 3 weeks on a phosphorimager screen).

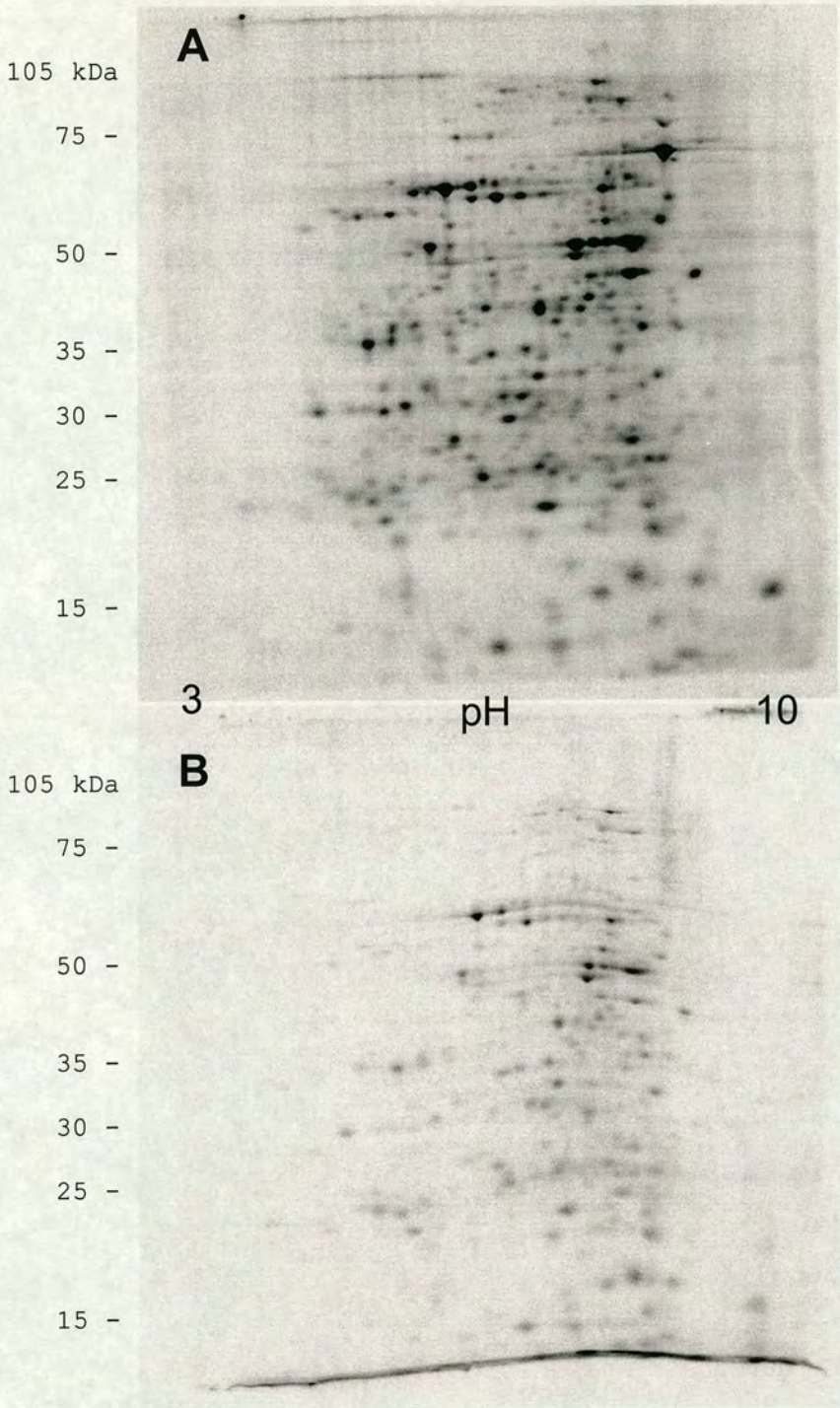


Fig. 6.3.3 2D gels of soluble proteins from MGM100 labelled with ^{35}S-methionine. MGM100 was grown in the presence of ara (A) or glu (B) and labelled with ^{35}S-methionine, the soluble protein fraction was then separated by 2D electrophoresis. These images show that the efficiency of labelling of newly synthesised proteins was greatly reduced in GroE depleted (B) cells.

6.4 Growth of GroE depleted cells in SCAA

SCAA –Met was formulated so that radiolabelling experiments in GroE depleted MGM100 could be performed; it was therefore necessary to determine if GroE depleted MGM100 acted in the same way in this medium as it did in LB. MGM100 was grown in appropriate medium containing arabinose until OD_{600} ~ 0.2 and then diluted 1:10 into fresh medium containing glucose. The results (Fig. 6.4.1) show that GroE depleted MGM100 grown in MM1 SCAA behaves like GroE depleted MGM100 growing in LB; the cells grow at a decreasing rate and eventually lyse. But GroE depleted MGM100 growing in MM1 SCAA –Met behave differently; the cells grow at a decreasing rate, but do not lyse (acting like MGM100 grown in LB, glu and DAP). This was an unexpected result, which meant that MM1 SCAA – Met was not suitable for use in radiolabelling of proteins in GroE depleted cells. Since MGM100 did not lyse in this medium there was the possibility that the results obtained from this medium would be different from results obtained from MGM100 grown in LB + glucose.

To determine if other amino acid deficient media caused unusual growth in GroE depleted MGM100, multiple versions of SCAA were made each of which lacked one amino acid. The growth curves of GroE depleted MGM100 growing in these media can be seen in figures 6.4.2 and 6.4.3. While growth rates and lysis times of MGM100 varied in these different media, lysis eventually occurred in all but three: SCAA –Leu, SCAA –Lys and SCAA –Met.

MGM100 had a very slow growth rate in the medium lacking isoleucine. In *E. coli*, K-12 strain, the isoleucine and valine biosynthetic pathways are parallel, with one enzyme catalysing both parallel reactions (Umbarger, 1996). One of the enzymes on the biosynthetic pathways (acetohydroxy acid synthase) is only derepressed when valine is limiting [the isozyme of acetohydroxy acid synthase that is insensitive to valine inhibition contains a frameshift mutation, and is therefore not active (Lawther et al, 1981)]. Therefore when isoleucine levels are low, but valine levels are high, the strain cannot synthesise isoleucine, and therefore grows at a very slow rate (if at all). It is important to stress that this is not a result of GroE depletion.

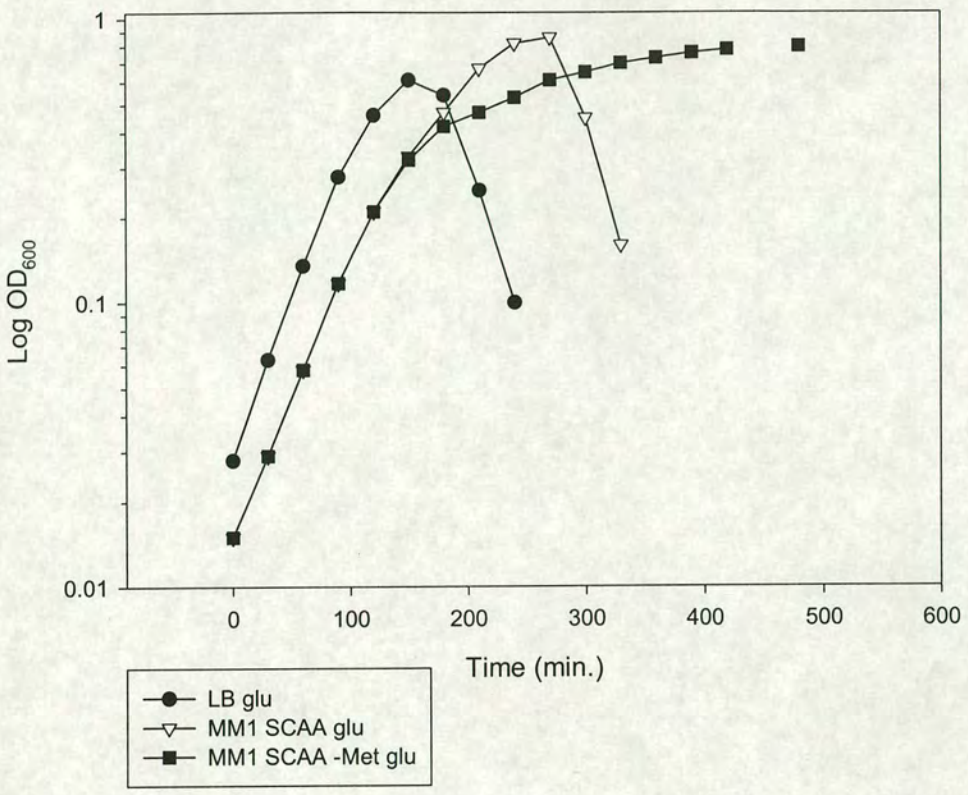


Fig. 6.4.1 Growth of MGM100 in various media. MGM100 was grown overnight in an appropriate medium, LB and ara, MM1 SCAA and ara or MM1 SCAA –Met and ara, at 37°C with constant agitation. The following day the culture was diluted 1:100 into fresh medium and grown at 37°C with constant agitation until the OD_{600} was ~ 0.2. 1 ml of culture was removed, centrifuged and the supernatant discarded. The cell pellet was then resuspended in 10 ml of prewarmed medium and glucose (this was time = 0). The cells were grown at 37°C with constant agitation and the OD_{600} was determined every 30 minutes. The cultures were diluted 1:10 when the OD_{600} reached ~ 0.2. For clarity, where the cultures have been diluted 1:10, the OD has been multiplied by 10 and plotted. All cultures were diluted 1:10 at 90 minutes.

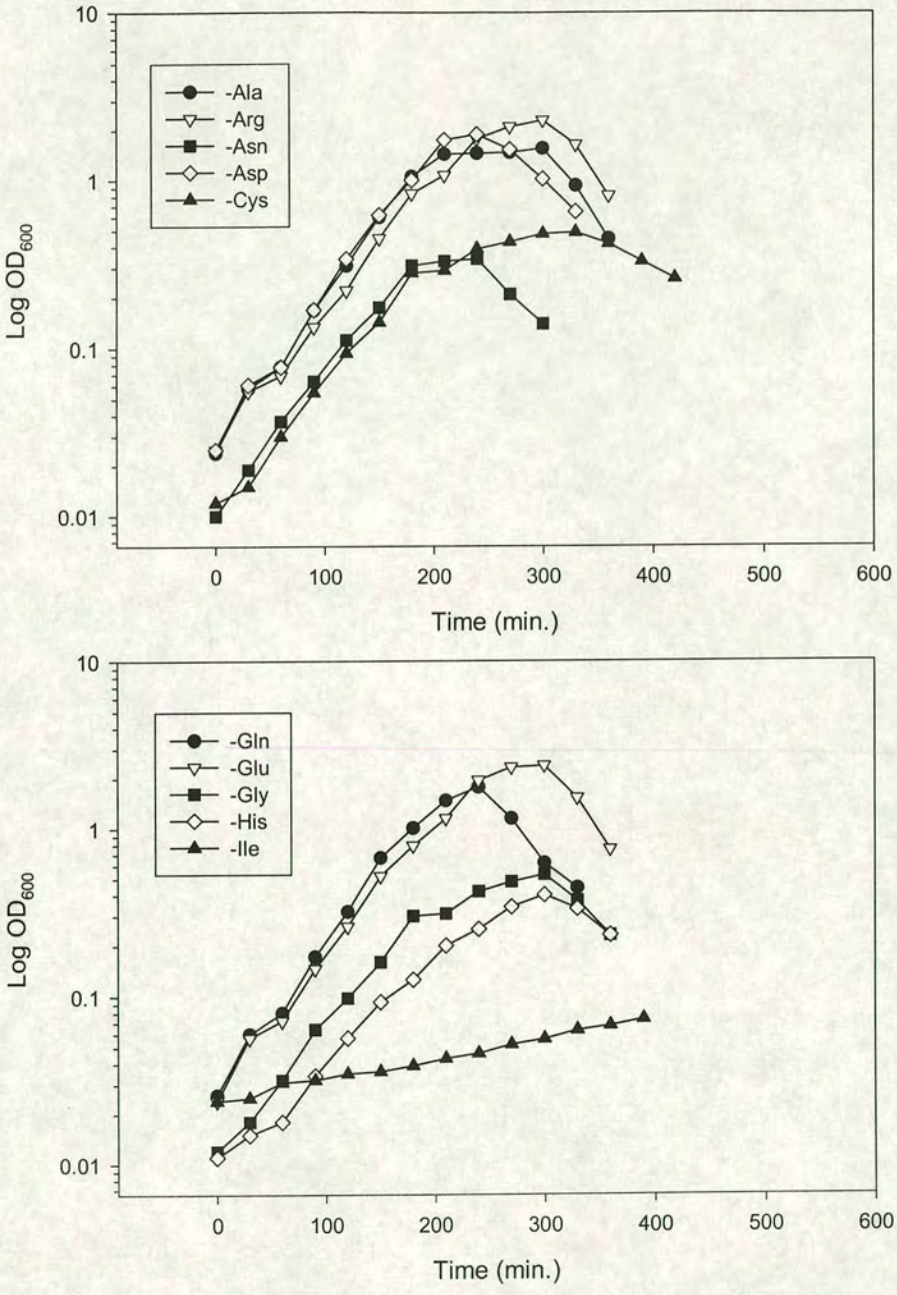


Fig. 6.4.2 Growth of MGM100 in various media. MGM100 was grown overnight in an appropriate medium (MM1 SCAA –various amino acids) and ara at 37°C with constant agitation. The following day the culture was diluted 1:100 into fresh medium and grown at 37°C with constant agitation until the OD_{600} was ~ 0.2. 1 ml of culture was removed, centrifuged and the supernatant discarded. The cell pellet was then resuspended in 10 ml of prewarmed medium and glucose (this was time = 0). The cells were grown at 37°C with constant agitation and the OD_{600} was determined every 30 minutes. The cultures were diluted 1:10 when the OD_{600} reached ~ 0.2. For clarity, where the cultures have been diluted 1:10, the OD has been multiplied by 10 and plotted.

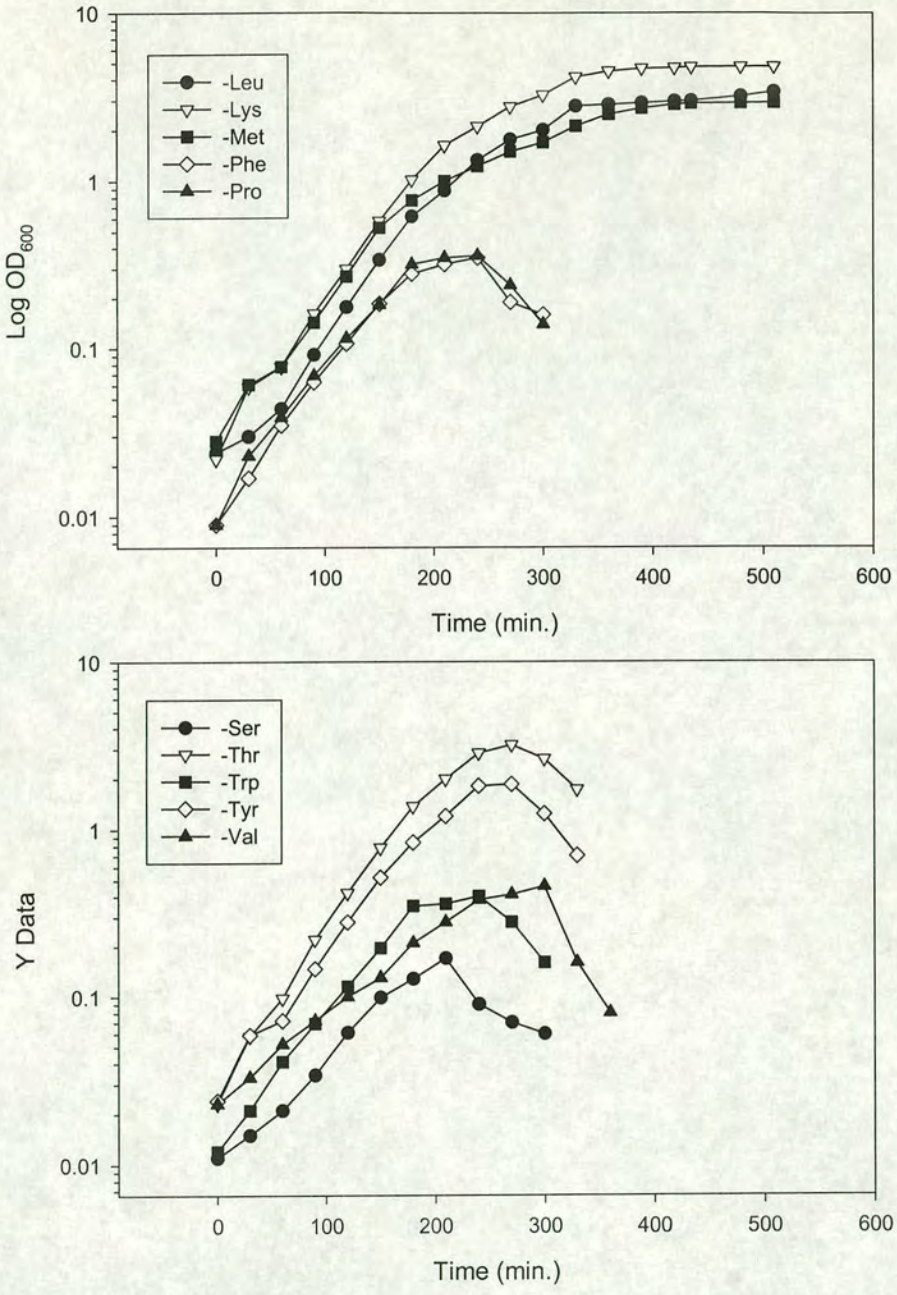


Fig. 6.4.3 Growth of MGM100 in various media. MGM100 was grown overnight in an appropriate medium (MM1 SCAA –various amino acids) and ara at 37°C with constant agitation. The following day the culture was diluted 1:100 into fresh medium and grown at 37°C with constant agitation until the OD_{600} was ~ 0.2. 1 ml of culture was removed, centrifuged and the supernatant discarded. The cell pellet was then resuspended in 10 ml of prewarmed medium and glucose (this was time = 0). The cells were grown at 37°C with constant agitation and the OD_{600} was determined every 30 minutes. The cultures were diluted 1:10 when the OD_{600} reached ~ 0.2. For clarity, where the cultures have been diluted 1:10, the OD has been multiplied by 10 and plotted.

6.5 MGM100 in MM1 SCAA –Lys

When GroE depleted MGM100 is grown in medium lacking lysine, the cells do not lyse (though they do eventually stop growing). In chapter 5 it is shown that expression from the *dapA* promoter increases dramatically during GroE depletion in lysine deficient medium. Also if GroE depleted cells have enough extra DapA before GroE depletion, they do not lyse (see chapter 3.9). Finally, it is also known that lysine inhibits DapA activity (Laber et al, 1992). From these facts it was hypothesised that during GroE depletion in lysine deficient medium, as DapA started to become inactive (due to GroE depletion), DAP levels decrease. Expression from the *dapA* promoter would then increase as DAP is required for the synthesis of both peptidoglycan and lysine. The increased synthesis of new DapA would raise the level of DapA activity (the cells would not be completely GroE depleted by this time, and therefore some percentage of newly synthesised DapA would be folded correctly and consequently active; also since lysine levels would be decreased, this DapA would be more active). This increased level of DapA activity (relative to DapA activity in MGM100 grown in medium containing lysine) would delay the process of lysis. The cells would then stop growing (due to GroE depletion) before lysis occurred, and hence there would be no lysis in GroE depleted cells growing in lysine deficient medium.

To determine if this hypothesis was correct a number of experiments were performed.

6.5.1 P_{dapA} *activity in lysine deficient medium*

MGM100λP_{dapA}-*lacZ* was grown in MM1 supplemented with either SCAA or SCAA –Lys. The cells were grown as described in section 6.4 in the presence of either ara or glu. During growth in these media samples were taken, and β-galactosidase activity (and hence P_{dapA} activity) was determined. The results of this experiment are shown in figure 6.5.1. These results indicate that in GroE depleted cells growing in lysine deficient medium (MM1 SCAA –Lys) *dapA* promoter activity is significantly higher than in GroE depleted cells growing in medium containing lysine. This would result in more DapA being made in GroE depleted cells growing in lysine deficient medium than in medium containing lysine.

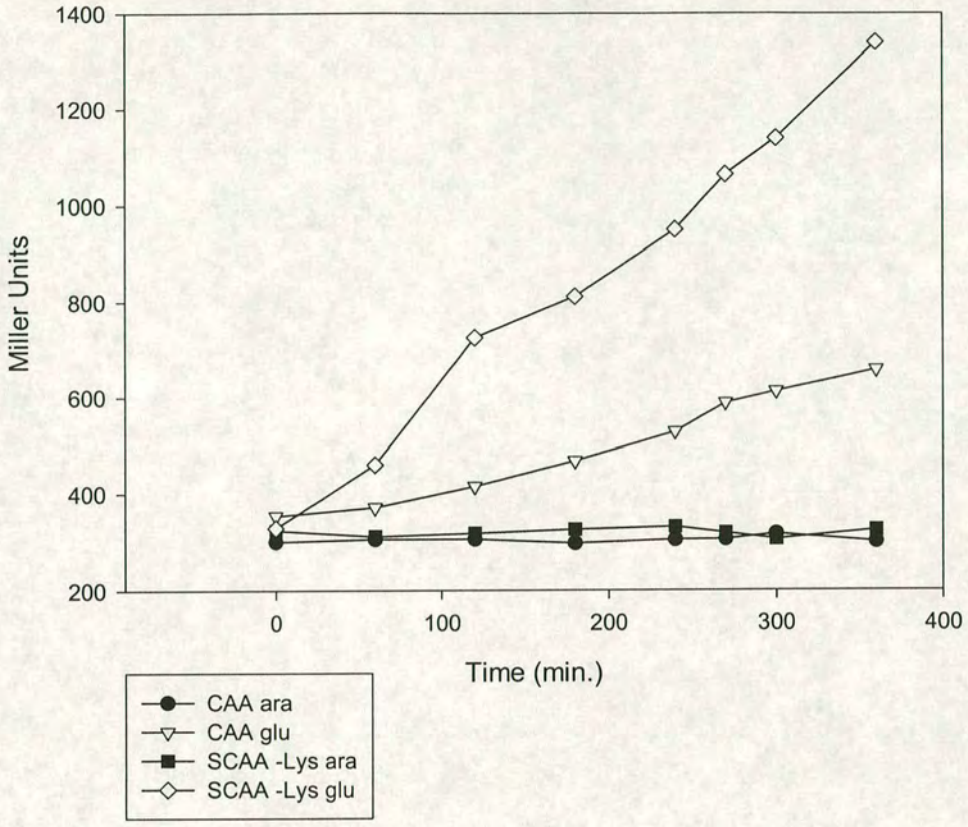


Fig 6.5.1 β-galactosidase activity in MGM100λP_{dapA}-*lacZ* growing in MM1 supplemented with CAA, or SCAA –Lys. MGM100λP_{dapA}-*lacZ* was grown overnight in an appropriate medium and ara at 37°C with constant agitation. The following day the culture was diluted 1:100 into fresh medium and grown at 37°C with constant agitation until the OD_{600} was ~ 0.2. 1 ml of culture was removed, centrifuged and the supernatant discarded. The cell pellet was then resuspended in 10 ml of prewarmed medium and ara or glu (this was time = 0). The cells were grown at 37°C with constant agitation and the β-galactosidase activity was determined at various times. These results show that β-galactosidase activity (and hence *dapA* promoter activity) increases as cells become GroE depleted, with *dapA* promoter activity being much higher in lysine deficient medium.

6.5.2 DapA activity in lysine deficient medium

An experiment was performed to determine if DapA activity was higher in GroE depleted cells growing in lysine deficient medium than in medium containing lysine. MGM100 was grown as described in section 6.5.1, except that instead of taking samples for β-galactosidase assays, samples were taken for DapA activity assays (as described in materials and methods). The results of these assays (Fig. 6.5.2) show that in $GroE^+$ cells DapA activity is constant (though it is slightly higher in lysine deficient medium, this may be because the cells have a lower internal level of lysine and therefore DapA is not inhibited to the same degree as a cell growing in medium containing lysine). But in $GroE^-$ cells DapA activity declines; though GroE depleted cells growing in lysine deficient medium have a higher level of DapA activity at later time points than GroE depleted cells growing in medium containing lysine. This is good evidence that the original hypothesis was correct and that GroE depleted cells in lysine deficient medium do not lyse because they have enough active DapA to prevent them from lysing until the cells stop growing (due to GroE depletion). Once the cells have stopped growing they cannot lyse (lysis is due to the cells not being able to make peptidoglycan quickly enough; if the cells have stopped growing, they no longer need to make new peptidoglycan).

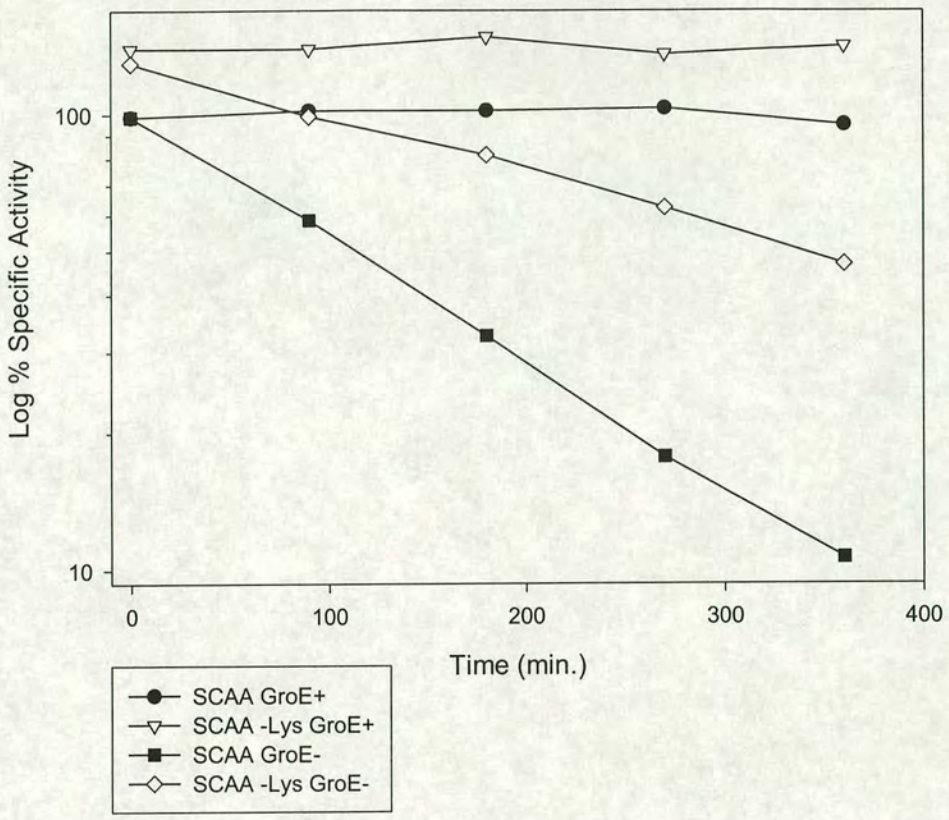


Fig 6.5.2 DapA specific activity in MGM100 under various conditions. MGM100 was grown overnight in either MM1 SCAA ara or MM1 SCAA –Lys ara at 37°C with constant agitation. The following day the culture was diluted 1:100 into fresh medium and grown at 37°C with constant agitation until the OD_{600} was ~ 0.2. 1 ml of culture was removed, centrifuged and the supernatant discarded. The cell pellet was then resuspended in 10 ml of prewarmed medium and ara ($GroE^+$) or glu ($GroE^-$) this was time 0. The cells were grown at 37°C with constant agitation and samples taken at various times. The DapA activity was determined per mg of total protein and plotted on the graph. The sample SCAA $GroE^+$ was set as 100% and all other sample activities are plotted relative to it.

6.6 MGM100 in MM1 SCAA –Met

GroE depleted MGM100 growing in MM1 supplemented with SCAA –Met acted very much like MGM100 growing in SCAA –Lys. The cells grew at a decreasing rate, but did not lyse (though they did stop growing). Unlike GroE depletion in lysine deficient medium, nothing was known about what happens at the intracellular level to GroE depleted cells growing in methionine deficient medium.

The biosynthesis of methionine in *E. coli* (Fig. 6.6.1) and *S.typhimurium* has been extensively reviewed (Cohen et al, 1987; Old et al, 1991; Saint-Girons et al, 1988; Flavin, 1975). Briefly, methionine has been classified as a member of the aspartate family of amino acids since the atoms of its four-carbon chain are derived from aspartic acid. The first three reactions are shared with the biosynthetic pathways of several other metabolites, including lysine, threonine and the branched chain amino acids. The other reactions of the pathway, which convert homoserine and methylenetetrahydrofolate to methionine, are uniquely concerned with methionine synthesis. The loss of any or all of the enzymes that catalyse them creates a nutritional requirement for methionine only.

Using two key pieces of information a hypothesis was formed to explain the observation the GroE depleted cells growing in methionine deficient medium do not lyse. The first piece of information is that during GroE depletion in MGM100 (growing in LB or minimal medium) there is a large induction of *metE* [Fig. 6.6.2, the *metE* band was identified by N-terminal sequencing (N. McLennan, pers. comm.)] and also *metC* (though this cannot be seen clearly on the gel images obtained) this has also been reported elsewhere (Horwich et al, 1993). The second piece of information is the observation that diaminopimelic acid can be replaced by cystathionine or lanthionine in the peptidoglycan of *E. coli* (Mengin-Lecreulx et al, 1994; Richaud et al, 1993). This is because cystathionine and lanthionine have a very similar structure to diaminopimelic acid (Fig. 6.6.3) and can therefore be incorporated into peptidoglycan by MurE when meso-DAP levels are low.

Using these two pieces of information the following hypothesis was formulated; in GroE depleted cells DapA activity decreases, consequently meso-DAP is not made and cannot be incorporated into peptidoglycan. New peptidoglycan therefore may utilise either cystathionine or lanthionine (or possibly both) to replace DAP. Since cystathionine is on the pathway to methionine biosynthesis, if it were diverted to peptidoglycan synthesis, internal levels of homocysteine and methionine would decrease. Expression of *metE* is regulated by levels of both homocysteine and of methionine (Greene, 1996; Wu et al, 1993); therefore as the levels of homocystine

and methionine decrease there is an increased expression of *metE*. Also, since the cells are growing in methionine deficient medium, and they are possibly not synthesising methionine, the cells stop growing (there may also be a problem with methionine uptake in GroE depleted cells, which would explain why MetE is also induced in medium containing methionine i.e. LB). If the cells stop growing before lysis occurs this may prevent lysis from occurring at all.

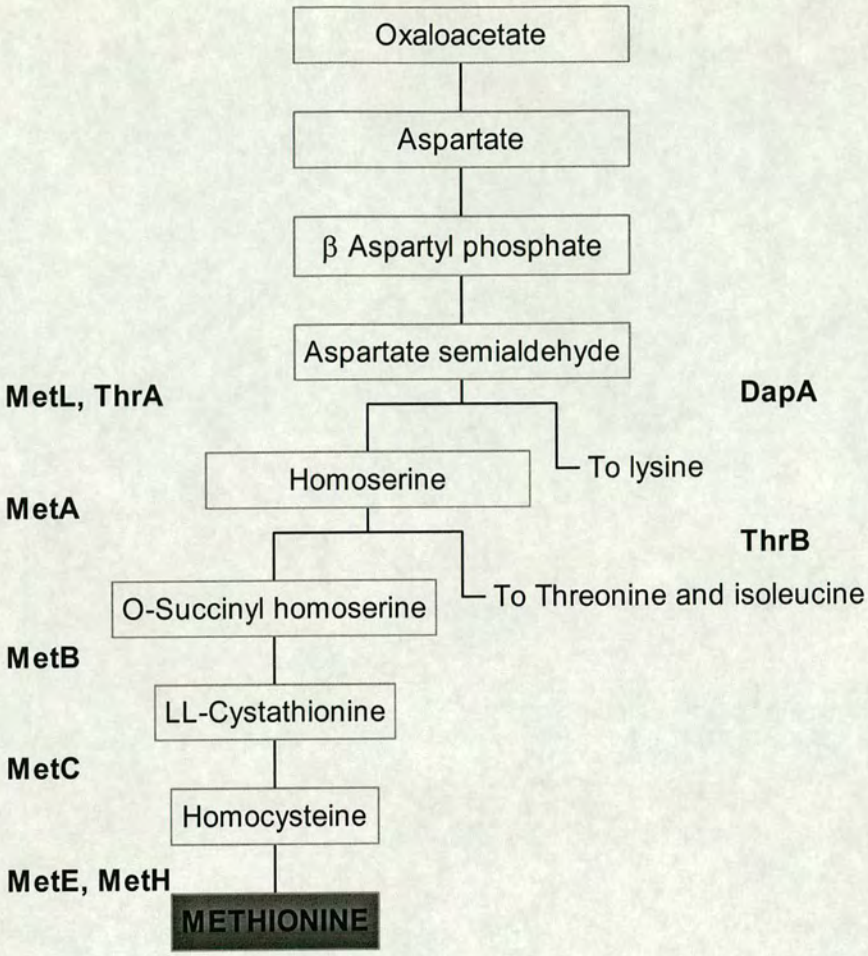


Fig 6.6.1 Biosynthesis of methionine in *E. coli*. The product of the *metL* gene is homoserine dehydrogenase II; *metA*: homoserine transsuccinylase; *metB*: cystathionine γ-synthase; *metC*: β-cystathionase; *metE* and *metH*: methionine synthase; *thrA*: homoserine dehydrogenase I; *thrB*: homoserine kinase; *dapA*: dihydropicolinate synthetase.

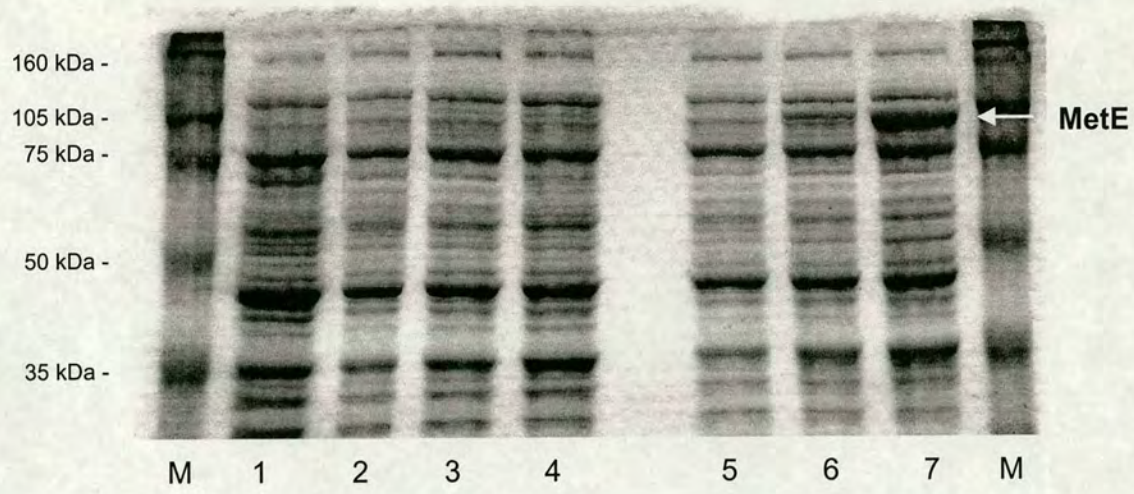


Fig. 6.6.2 Coomassie blue stained SDS gel of cell lysate of MGM100 grown in various conditions. Lanes; M: molecular weight markers; 1: $GroE^{+/-}$ sample, 0 minutes; 2: $GroE^{+}$ sample, 45 minutes; 3: $GroE^{+}$ sample, 90 minutes; 4: $GroE^{+}$ sample, 135 minutes; 5: $GroE^{-}$ sample, 45 minutes; 6: $GroE^{-}$ sample, 90 minutes; 7: $GroE^{-}$ sample, 135 minutes. MGM100 was grown overnight in LB and ara at 37°C with constant agitation. The following day the culture was diluted 1:100 into fresh medium and grown at 37°C with constant agitation until the OD_{600} was ~ 0.2. 1 ml of culture was removed, centrifuged and the supernatant discarded. The cell pellet was then resuspended in 10 ml of prewarmed medium and ara ($GroE^{+}$) or glu ($GroE^{-}$) this was time 0. The cells were grown at 37°C with constant agitation and samples taken at various times to run on an SDS polyacrylamide gel. In the $GroE^{-}$ samples (especially lane 7) MetE (84.6 kDa) is clearly induced.

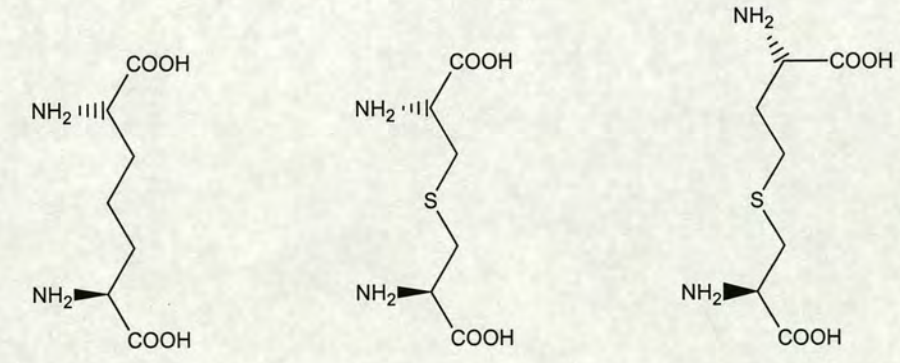


Fig. 6.6.3 Structural relationships between stereoisomers of cysteine thioethers and of diaminopimelate.

To determine if the hypothesis was correct, DAP was added to GroE depleted cells. If DAP is present then it should be used preferentially instead of cystathionine for peptidoglycan synthesis. Under these conditions cystathionine should then only be used for methionine synthesis, and hence methionine levels should not decrease. When DAP is added to GroE depleted cells, however, MetE is still induced (Fig. 6.6.4).

This observation suggests that MetE induction may be due to GroE depletion affecting some aspect of methionine synthesis directly, rather than indirectly (i.e. the DapA/cystathionine hypothesis). To determine if the induction was caused by GroE depletion or DAP depletion, the strain MGM100Δ*dapA* was grown in osmotically protected medium and ara (under these conditions the strain will be DAP depleted but will remain $GroE^+$), and samples were run on an SDS-PAGE gel. The results (Fig 6.6.5) show that when MGM100 is DAP depleted but has wild type levels of GroE, there is no induction of MetE. Also when homoserine (100μg ml^{-1}) was added to GroE depleted cells, MetE was still induced (Fig. 6.6.5). This data suggests that GroE depletion is causing MetE induction, but not because cystathionine is being diverted to make peptidoglycan, and since even the addition of homoserine (the immediate methionine precursor) did not affect MetE induction, there is the possibility that methionine levels in GroE depleted cells are not reduced.

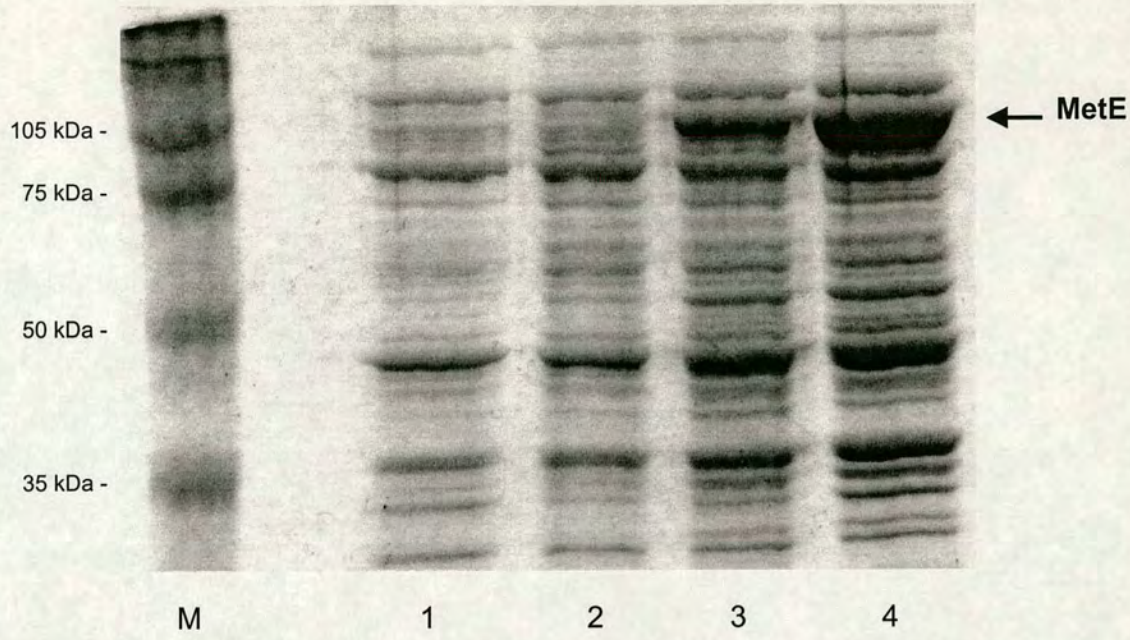


Fig. 6.6.4 Coomassie blue stained SDS gel of cell lysate of MGM100 grown in LB and DAP. Lanes; M: molecular weight markers; 1: GroE$^-$ sample, 0 minutes; 2: GroE$^-$ sample, 45 minutes; 3: GroE$^-$ sample, 90 minutes; 4: GroE$^-$ sample, 135 minutes. MGM100 was grown overnight in LB and ara at 37°C with constant agitation. The following day the culture was diluted 1:100 into fresh medium and grown at 37°C with constant agitation until the OD_{600} was ~ 0.2. 1 ml of culture was removed, centrifuged and the supernatant discarded. The cell pellet was then resuspended in 10 ml of prewarmed medium, glu and DAP (GroE$^-$) this was time 0. The cells were grown at 37°C with constant agitation and samples taken at various times to run on an SDS polyacrylamide gel. In the GroE$^-$ samples (especially lanes 3 and 4) MetE (84.6 kDa) is clearly induced.

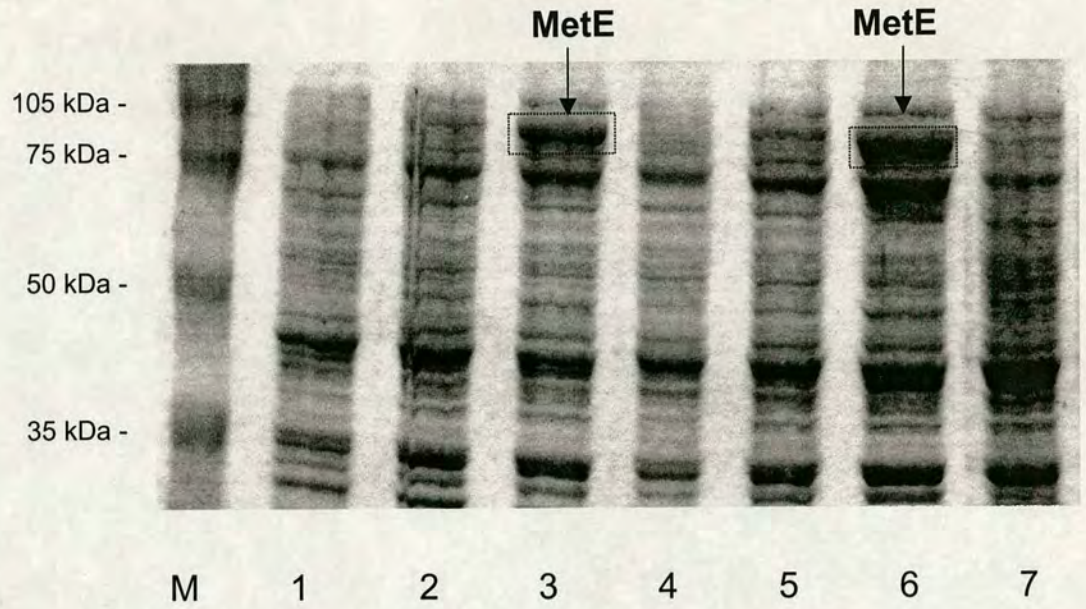


Fig. 6.6.5 Coomassie blue stained SDS gel of cell lysate of MGM100 and MGM100Δ*dapA* grown under various conditions. Lanes; M: molecular weight markers; 1-3: MGM100 GroE$^-$ sample, 45, 90 and 135 minutes respectively; 4-6: MGM100 GroE$^-$ sample + homoserine, 45, 90 and 135 minutes respectively; 7: MGM100Δ*dapA* grown in osmotically protected MM1 and ara, 60 minutes. MGM100 or MGM100Δ*dapA* was grown in an appropriate medium at 37°C with constant agitation. The following day the culture was diluted 1:100 into fresh medium and grown at 37°C with constant agitation until the OD_{600} was ~ 0.2. 1 ml of culture was removed, centrifuged and the supernatant discarded. The cell pellet was then resuspended in 10 ml of prewarmed medium, glu and DAP (GroE$^-$) or in the case of MGM100Δ*dapA* the pellet was resuspended in osmotically protected MM1 and ara (this was time 0). The cells were grown at 37°C with constant agitation and samples taken at various times to run on an SDS polyacrylamide gel. In the GroE$^-$ samples (especially lanes 3 and 6) MetE (84.6 kDa) is clearly induced. MetE is also induced homoserine is present. This suggests that methionine feedback activation by methionine is not causing MetE induction. In the GroE$^+$, DAP depleted sample (lane 7) there is no induction of MetE, suggesting that MetE induction is caused by GroE depletion, and not cystathionine depletion.

6.7 Summary and discussion

In this chapter work has been described that attempts to explain unusual phenomena observed when GroE depleted cells are grown in media deficient in individual amino acids.

Initially a supplement was created that contained the same level of amino acids as a 100 mg ml^{-1} solution of casamino acid. The amount of each amino acid was determined from previously published data concerning the composition of casein from bovine milk (Swaisgood, 1992). The medium was originally intended to be used during radiolabelling of MGM100 using ^{35}S-methionine. Because the cells could be grown in the medium (as they had a growth rate similar to that in LB) and also labelled in it as well (without the need for changing media) it was thought suitable. But when MGM100 was grown in this medium under conditions that caused GroE depletion, the cells no longer behaved like they would have if growing in LB.

It was found that when growing GroE depleted MGM100 in media deficient in either leucine, lysine or methionine, the cells did not lyse. In this chapter experiments are described which attempt to explain why GroE depleted cells do not lyse in lysine or methionine deficient media (due to time constraints no experiments were performed to try and explain the phenomena in leucine deficient medium).

While GroE depletion in lysine deficient medium had not been studied previously in this work, some observations regarding MGM100 in lysine deficient medium had been made (chapter 5). From these observations the following hypothesis was proposed: during GroE depletion, DapA activity decreases (chapter 3); as DapA activity decreases, DAP levels decrease. Since the cells are growing in lysine deficient medium, DAP must be used for both the synthesis of lysine and peptidoglycan, consequently DAP levels decrease faster than in equivalent cells growing in a medium containing lysine. As DAP levels decrease, *dapA* promoter activity increases (chapter 5), and DapA is made. While most of this DapA will be inactive, some fraction will retain activity (chapter 3) and consequently, DapA activity in GroE depleted cells will be higher than in equivalent cells growing in medium containing lysine (also there will be no inhibition of activity by lysine). As DapA activity is higher, this will prevent the cells from lysing long enough so that they stop growing (due to some other aspect of GroE depletion; chapter 1). Once the cells have stopped growing, lysis cannot occur.

To determine if this hypothesis was correct, *dapA* promoter activity was measured in MGM100 growing in medium with or without lysine. The promoter activity was higher in GroE depleted cells growing in lysine deficient medium than

in cell grown in medium containing lysine. The next step was to measure DapA activity in these two media. It was observed that DapA activity was higher in GroE depleted cells growing in medium deficient in lysine than in equivalent cells growing in medium containing lysine, at all time points (at 360 minutes after GroE depletion, MGM100 in MM1 SCAA had 10.6% wild type DapA activity, while MGM100 in MM1 SCAA –Lys had 46.3% wild type DapA activity). Since MGM100 in MM1 SCAA -Lys stopped growing at ~ 400 minutes after GroE depletion and had not lysed by this time, this suggested that the hypothesis might be correct.

A hypothesis was also created to try and explain why GroE depleted cells growing in methionine deficient medium did not lyse. This hypothesis was also, in some aspect, related to the decrease of DapA activity in GroE depleted cells. It was hypothesised that decreased DAP levels (due to the lack of DapA activity in GroE depleted cells) caused cystathionine to be incorporated into peptidoglycan in its place. If cystathionine was being incorporated into peptidoglycan, then there would be decreased levels and consequently, methionine biosynthesis would be decreased. Since both methionine and its immediate precursor, homoserine, regulate expression of MetE, this could explain the induction of MetE in GroE depleted cells. If the cells were not making methionine, and methionine was not present in the medium, then the cells would stop growing. Also, it was observed that methionine uptake was greatly reduced in GroE depleted cells; therefore cells growing in medium containing methionine (i.e. LB) may still have reduced internal levels of methionine when GroE is depleted. If this occurred before lysis due to GroE depletion, then the cells would not lyse.

To determine if this hypothesis was correct, DAP was added to GroE depleted cells growing in methionine deficient medium; if exogenous DAP were being incorporated into peptidoglycan, then cystathionine would not need to be diverted from methionine synthesis, and consequently MetE should not be induced. This did not prove to be the case; when exogenous DAP was added, MetE was still induced.

In an attempt to simulate the state of DapA in MGM100 without the cell being GroE depleted, the strain MGM100Δ*dapA* was used. This strain has no DapA activity (like GroE depleted MGM100) but remains $GroE^{+}$. This strain was grown in osmotically protected medium (without exogenous DAP) to see if MetE induction was caused by DAP depletion or GroE depletion. The strain did not show an increased expression of MetE. This data suggests that GroE depletion was causing MetE induction through some other pathway that did not involve DapA or DAP.

When homoserine was added to GroE depleted cells, MetE was still induced. This suggests that MetE itself may be affected by GroE depletion (if MetE is inactive

in GroE depleted cells, it will not be able to catalyse the transformation of homoserine to methionine, consequently methionine will not be made, and MetE will be expressed). In all of the above experiments the level of denatured MetE protein has been observed, but not the activity of MetE. Another possibility is that MetE induction is not being caused by decreased methionine levels, but by some other aspect of GroE depletion. Horwich (Horwich et al, 1993) suggests that MetE induction represents a mechanism of regulation of translational initiation in GroE depleted cells. Either way, more work is obviously necessary before it will be possible to explain either MetE induction in GroE depleted cells, or why GroE depleted cells do not lyse in methionine deficient medium.

Chapter VII

References

Alexander, F. W., Sandmeier, E., Mehta, P. K., and Christen, P. 1994. Evolutionary relationships among pyridoxal-5'-phosphate-dependent enzymes. Regio-specific alpha, beta and gamma families. *Eur.J.Biochem.* **219**: 953-960.

Amann, E., Brosius, J., and Ptashne, M. 1983. Vectors bearing a hybrid *trp-lac* promoter useful for regulated expression of cloned genes in *Escherichia coli. Gene* **25**: 167-178.

Andersen, A. B. and Hansen, E. B. 1993. Cloning of the *lysA* gene from *Mycobacterium tuberculosis. Gene* **124**: 105-109.

Anderson, W. F. and Steitz, T. 1975. Structure of yeast hexokinase. IV. Low-resoultion structure of enzyme-substrate complexes revealing negative co-operativity and allosteric interactions. *J.Mol.Biol.* **92**: 279-287.

Anfinsen, C. B. 1973. Principles that govern the folding of protein chains. *Science* **181**: 223-230.

Ang, D., Keppel, F., Klein, G., Richardson, A., and Georgopoulos, C. 2000. Genetic analysis of bacteriophage-encoded cochaperonins. *Annu.Rev.Genet.* **34**: 439-456.

Antia, M., Hoare, D. S., and Work, E. 1957. The stereoisomers of αε-diaminopimelic acid 3. Properties and distributions of diaminopimelic acid racemase, an enzyme causing interconversion of the LL and *meso* isomers. *Biochem.J.* **65**: 448-459.

Asada, Y., Tanizawa, K., Sawada, S., Suzuki, T., Misono, H., and Soda, K. 1981. Stereochemistry of meso-alpha,epsilon-diaminopimelate decarboxylase reaction: the first evidence for pyridoxal 5'-phosphate dependent decarboxylation with inversion of configuration. *Biochemistry* **20**: 6881-6886.

Bachmann, B. J. 1983. Linkage map of *Escherichia coli* K-12, edition 7. *Microbiol.Rev.* **47**: 180-230.

Baldwin, R. L. 1994. Finding protein intermediates in protein folding. *Bioessays* **16**: 207-210.

Baneyx, F. and Gatenby, A. A. 1992. A mutation in GroEL interferes with protein folding by reducing the rate of discharge of sequestered polypeptides. *J.Biol.Chem.* **267**: 11637-11644.

Barbier, B., Charlier, M., and Maurizot, J. C. 1984. Photochemical cross-linking of *lac* repressor to nonoperator 5-bromouracil-substituted DNA. *Biochemistry* **23**: 2933-2939.

Barraclough, R. and Ellis, R. J. 1980. Assembly of newly synthesised large subunits into ribulose bisphosphate carboxylase in isolated intact pea chloroplasts. *Biochim.Biophys.Acta.* **608**: 19-31.

Beaman, T. W., Binder, D. A., Blanchard, J. S., and Roderick, S. L. 1997. Three-dimensional structure of tetrahydrodipicolinate N-succinyltransferase. *Biochemistry* **36**: 489-494.

Berges, D. A., DeWolf, W. E., Dunn, G. L., Grappel, S. F., Newmann, S. J., Taggart, J. J., and Gilvarg, C. 1986. Studies on the active site of succinyl-CoA:tetrahydrodipicolinate N-succinyltransferase. Characterization using analogs of tetrahydrodipicolinate. *J.Biol.Chem.* **26129**: 6160-6167.

Berges, D. A., DeWolf, W. E., Dunn, G. L., Grappel, S. F., Newmann, S. J., Taggart, J. J., and Gilvarg, C. 1986. Peptides of 2-aminopimelic acid: antibacterial agents that inhibit diaminopimelic acid biosynthesis. *J.Med.Chem.* **29**: 89-95.

Biellmann, J. F., Eid, P., Hirth, C., and Jornvall, H. 1980. Aspartate-β-semialdehyde dehydrogenase from *Escherichia coli*: Purification and general properties. *Eur.J.Biochem.* **104**: 53-58.

Black, S. and Wright, N. G. 1955. Aspartic β-semialdehyde dehydrogenase and aspartic β-semialdehyde. *J.Biol.Chem.* **213**: 39-50.

Blickling, S., Renner, C., Laber, B., Pohlenz, H. D., Holak, T. A., and Huber, R. 1997. Reaction mechanism of Escherichia coli dihydrodipicolinate synthase investigated by X-ray crystallography and NMR spectroscopy. *Biochemistry* **36**: 24-33.

Bochkareva, E. S., Lissin, N. M., and Girshovich, A. S. 1988. Transient association of newly synthesised unfolded proteins with the heat shock GroEL protein. *Nature* **336**: 254-257.

Bolivar, F. 1978. Construction and characterisation of new cloning vehicles III. Derivatives of plasmid pBR322 carrying unique *Eco*RI sites for selection of *Eco*RI generated recombinant DNA molecules. *Gene* **4**: 121-136.

Bonnassie, S., Oreglia, J., and Sicard, A. M. 1990. Nucleotide sequence of the *dapA* gene from *Corynebacterium glutamicum*. *Nucl.Acids Res.* **18**: 6421.

Borowiec, A. and Gralla, J. D. 1987. All the elements of the *lac* ps promoter mediate its transcriptional response to DNA supercoiling. *J.Mol.Biol.* **195** : 89-97.

Borowiec, A., Zhang, L., Sasse-Dwight, S., and Gralla, J. D. 1987. DNA supercoiling promotes formation of a bent repression loop in *lac* DNA. *J.Mol.Biol.* **196**: 101-111.

Bouvier, J., Richaud, F. C., Richaud, F., Patte, J. C., and Stragier, P. 1984. Nucleotide sequence and expression of the *Escherichia coli dapB* gene. *J.Biol.Chem.* **259**: 14829-14834.

Boy, E. and Patte, J. C. 1972. Multivalent repression of aspartic semialdehyde dehydrogenase in *Escherichia coli* K-12. *J.Bacteriol.* **112**: 84-92.

Boy, E., Reinisch, F., Richaud, C., and Patte, J. C. 1976. Role of lysys-tRNA in the regulation of lysine biosynthesis in *Escherichia coli* K12. *Biochimie* **58**: 213-218.

Boy, E., Borne, F., and Patte, J. C. 1978. Effect if mutations affecting lysyl-$tRNA_{lys}$ on the regulation of lysine biosynthesis in *Escherichia coli*. *Mol.Gen.Genet.* **159**: 33-38.

Boy, E., Richaud, C., and Patte, J. C. 1979. Multiple regulation of DAP-decarboxylase synthesis in *Escherichia coli* K12. *FEMS Microbiol.Letts.* **5**: 287-290.

Boy, E., Borne, F., and Patte, J. C. 1979. Isolation and identification of mutants constitutive for aspartokinase III synthesis in *Escherichia coli* K12. *Biochimie* **61**: 1151-1160.

Boy, E. and Patte, J. C. 1979. Role of glucose 6-phosphate in the regulation of aspartate semialdehyde dehydrogenase in *Escherichia coli*. *FEMS Microbiol.Letts.* **6**: 189-192.

Boyen, A., Charlier, D., Charlier, J., Sakanyan, V., Mett, I., and Glansdorff, N. 1992. Acetylornithine deacetylase, succinyldiaminopimelate desuccinylase and carboxypeptidase G2 are evolutionarily related. *Gene* **116**: 1-6.

Bradford, M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal.Biochem.* **72**: 248-254.

Braig, K., Furuya, F., Hainfeld, J., and Horwich, A. L. 1993. The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* **371**: 578-586.

Brazil, B. T., Cleland, J. L., McDowell, R. S., Skelton, N. J., Paris, K., and Horowitz, P. M. 1997. Model peptide studies demonstrate that amphipathic secondary structures can be recognised by the chaperonin GroEL. *J.Biol.Chem.* **272**: 5105-5111.

Broquist, H. P., Stiffey, A. V., and Albrecht, A. M. 1961. Biosynthesis of lysine from α-ketoadipic acid and α-aminoadipic acid in yeast. *Appl.Microbiol.* **9**: 1.

Buchanan, R. L. and Gralla, J. D. 1987. Factor interactions at simian virus 40 GC-box promoter elements in intact nuclei. *Mol.Cell.Biol.* **7**: 1554-1558.

Buchner, J., Schmidt, M., Fuchs, M., Jaenicke, R, Rudolph, R, Schmid, F. X., and Kiefhaber, T. 1991. GroE facilitates refolding of citrate synthase by suppressing aggregation. *Biochemistry* **30**: 1586-1591.

Buckle, A. M., Zahn, R., and Fersht, A. R. 1997. A structural model for GroEL-polypeptide recognition. *Proc.Natl.Acad.Sci.U.S.A.* **94**: 3571-3575.

Bukhari, A. I. and Taylor, A. L. 1971. Genetic analysis of diaminopimelic acid and lysine-requiring mutants of *Escherichia coli*. *J.Bacteriol.* **105**: 988-998.

Bult, C. J., White, O., Olsen, G. J., Fleischamnn, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J. F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L., Geoghagen, N. S., and Venter, J. C. 1996. Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii. *Science* **273**: 1058-1073.

Butour, J. L., Felenbok, B., and Patte, J. C. 1974. Synthesis of dihydropicolinate synthetase in *Escherichia coli* K-12. *Ann.Microbiol.(Paris)* **125**: 459-462.

Cardineau, C. A. and Curtiss III, R. 1987. Nucleotide sequence of the *asd* gene of *Streptococcus mutans*. Identification of the promoter region and evidence for attenuator-like sequences preceeding the structural gene. *J.Biol.Chem.* **262**: 3344-3353.

Carrillo, N., Ceccarelli, E. A., Krapp, A. R., Boggio, S., Ferreyra, R. G., and Viale, A. M. 1992. Assembly of plant ferredoxin $NADP^+$ oxidoreductase in *Escherichia coli* requires GroE molecular chaperones. *J.Biol.Chem.* **267**: 15537-15541.

Cassan, M., Ronceray, J., and Patte, J. C. 1983. Nucleotide sequence of the promoter region of the *E. coli lysC* gene. *Nucl.Acids Res.* **11**: 6157-6166.

Chandrasekhar, G. N., Tilly, K., Woolford, C., Hendrix, R. W., and Georgopoulos, C. 1986. Purification and properties of the groES morphogenetic protein of *Escherichia coli*. *J.Biol.Chem.* **261**: 12414-12419.

Chatellier, J., Buckel, A. M., and Fersht, A. R. 1999. GroEL recognises sequential and non-sequential linear structural motifs compatible with extended beta strands and alpha helices. *J.Mol.Biol.* **292**: 163-172.

Chatellier, J., Hill, F., Foster, N. W., Goloubinoff, P., and Fersht, A. R. 2000. From minichaperone to GroEL 3: Properties of an active single-ring mutant of GroEL. *J.Mol.Biol.* **304**: 897-910.

Chatellier, J., Hill, F., and Fersht, A. R. 2000. From minichaperone to GroEL 2: importance of avidity of the multisite ring structure. *J.Mol.Biol.* **304**: 883-896.

Chen, L. and Sigler, P. B. 1999. The crystal structure of a GroEL/peptide complex: plasticity as a basis for substrate diveristy. *Cell* **99**: 757-768.

Chen, N. Y., Jiang, S. Q., Klein, D. A., and Paulus, H. 1993. Organisation and nucleotide sequence of the *Bacillus subtilis* diaminopimelate operon, a cluster of genes encoding the first three enzymes of diaminopimelate synthesis and dipicolinate synthesis. *J.Biol.Chem.* **262**: 8787-8798.

Chen, S., Roseman, A. M., Hunter, A. S., Wood, S. P., Burston, S. G., Ranson, N., Clarke, A. R., and Saibil, H. R. 1994. Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. *Nature* **371**: 261-264.

Cheshire, R. M. and Miflin, B. J. 1975. The control of lysine biosynthesis in maize. *Phytochem.* **14**: 695-698.

Christensen, H. and Pain, R. H. 1991. Molten globule intermediates and protein folding. *Eur.Biophys.* **19**: 221-229.

Church, G. M. and Gilbert, W. 1984. Genomic sequencing. *Proc.Natl.Acad.Sci.U.S.A.* **81**: 1991-1995.

Cohen, G. N., Patte, J. C., Truffa-Bachi, P., Sawas, C., and Doudoroff, M. 1965 *Mécanismes de régultaion des activités cellulaires chez les microrganismes.* Centre National del la Recherche Scientifique.

Cohen, G. N. 1985. Aspartate kinases I, II and III from *Escherichia coli*. *Methods Enzymol.* **113**: 596-599.

Cohen, G. N. and Saint-Girons, I. 1987 *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology.* American Society for Microbiology.

Cooper, A. 1999 *Protein: A Comprehensive Treatise* JAI Press Inc.

Coppo, A., Manzi, A., Pulitzer, J. F., and Takahashi, H. 1973. Abortive bacteriophage T4 head assembly in mutants of *Escherichia coli. J.Mol.Biol.* **76**: 61-87.

Corrales, F. J. and Fersht, A. R. 1996. Kinetic significance of $GroEL_{14}.(GroES_7)_2$ complexes in molecular chaperone activity. *Fold.Des.* **1**: 265-273.

Cox, R. J., Sherwin, W. A., Lam, L. K. P., and Vederas, J. C. 1996. Synthesis and Evaluation of Novel Substrates and Inhibitors of N-Succinyl-LL-diaminopimelate Aminotransferase (DAP-AT) from *Escherichia coli. J.Am.Chem.Soc.* **118** : 7449-7460.

Coyle, J. E., Jaeger, J., Radford, S. E., and Robinson, C. V. 1997. Structural and mechanistic consequences of polypeptide binding by GroEL. *Folding & Design* **2**: R93-R104.

Coyle, J. E., Texter, F. L., Ashcroft, A. E., Masselso, D., Robinson, C. V., and Radford, S. E. 1999. GroEL accelerates the refolding of hen lysozyme without changing its folding mechanism. *Nature Structural Biology* **6**: 683-690.

Crameri, A., Whitehorn, E. A., Tate, E., and Stemmer, W. P. C. 1996. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nature Biotechnol.* **14**: 315-319.

Creighton, T. E. 1990. Protein folding. *Biochem.J.* **270**: 1-16.

Cummins, C. S. and Harris, H. 1956. The chemical composition of the cell wall in some gram positive bacteria and its possible value as a taxonomic character. *J.Gen.Microbiol.* **14** : 583-600.

Daucé-Lereverend, B., Boitel, M., Deschamps, A. M., Lebeault, J. M., Sano, K., Takinami, K., and Patte, J. C. 1982. Improvement of *E. coli* strains overproducing lysine using recombinant DNA techniques. *Eur.J.Appl.Microbiol.Biotechnol.* **15**: 227-231.

Davis, B. D. 1952. Biosynthetic interrelations of diaminopimelic acid, lysine and threonine in mutants of *E. coli. Nature* **169**: 534.

De Wit, T. F. R., Bekelie, S>, Osland, A., Miko, T. L., Hermans, P. W. M., van Soolingen, P., Drijfhout, J. W., Schoeningh, R., Janson, A. A. M., and Thole, J. E. R. 1992. Mycobacteria contain two *groEL* genes: the seconds *Mycobacterium leprae groEL* gene is arranged in an operon with *groES*. *Mol.Microbiol.* **6**: 1995-2007.

Dereppe, C., Bold, G., Ghisalba, O., Ebert, E., and Schär, H. P. 1992. Purification and characterisation of dihydropicolinate synthetase from pea. *Plant Physiol.* **98**: 813-821.

Dessauer, C. W. and Bartlett, S. G. 1994. Identification of a chaperonin binding site in a chloroplast precursor protein. *J.Biol.Chem.* **269**: 19766-19776.

Dewey, D. C. and Work, E. 1952. Diaminopimelic acid decarboxylase. *Nature* **169**: 533-534.

Dionisi, H. M., Checa, S. K., Krapp, A. R., Arakaki, A. K., Ceccarelli, E. A., Carrillo, N., and Viale, A. M. 1998. Cooperation of the DnaK and GroE chaperone systems in the folding pathway of plant ferredoxin-NADP+ reductase expressed in *Escherichia coli*. *Eur.J.Biochem.* **251**: 724-728.

Dolan, K. M. and Greenberg, E. P. 1992. Evidence that GroEL, not σ32, is involved in transcriptional regulation of the *Vibrio fuscheri* luminescence genes in *Escherichia coli*. *J.Bacteriol.* **174**: 5132-5135.

Egan, S. E., Fleige, R., Tong, S., Shibata, A., Wolf, R. E., and Conway, T. 1992. Molecular characterisation of the Entner-Doudoroff pathway in *Escherichia coli*: sequence analysis and localisation of promoter from the *edd-eda* operon. *J.Bacteriol.* **174**: 4638-4646.

Eismann, E. R. and Muller-Hill, B. 1990. *lac* repressor forms stable loops in vitro with supercoiled wild-type *lac* DNA containing all three natural *lac* operators. *J.Mol.Biol.* **213**: 763-775.

Ellis, R. J. 1990. Molecular chaperones: the plant connection. *Science* **250**: 954-959.

Ellis, R. J. and van der Vies, S. M. 1991. Molecular chaperones. *Annu.Rev.Biochem.* **60**: 321-347.

Ellis, R. J. 1993. The general concept of molecular chaperones. *Phil.Trans.R.Soc.Lond.* **339**: 257-261.

Ellis, R. J. 1996 *The chaperonins* Academic Press.

Ewalt, K. L., Hendrick, J. P., Houry, W. A., and Hartl, F. U. 1997. *In vivo* observation of polypeptide flux through the bacterial chaperonin system. *Cell* **90**: 491-500.

Farber, G. and Petsko, G. 1990. The evolution of alpha/beta barrel enzymes. *Trends Biochem.Sci.* **15**: 228-234.

Farkas, W. and Gilvarg, C. 1965. The reduction step in diaminopimelic acid biosynthesis. *J.Biol.Chem.* **240**: 4717-4722.

Fenton, W. A., Kashi, Y., Furtak, K., and Horwich, A. L. 1994. Residues in chaperonin GroEL required for polypeptide binding and release. *Nature* **371**: 614-619.

Ferentz, A. E. and Wagner, G. 2000. NMR spectroscopy: a multifaceted approach to macromolecular structure. *Q.Rev.Biophys.* **33**: 29-65.

Fischer, H. M., Babst, M., Kaspar, T., Acuna, G., Arigoni, F., and Hennecke, H. 1993. One member of a *groESL*-like chaperonin multigene family in *Bradyrhizobium japonicum* is co-regulated with symbiotic nitrogen fixation genes. *EMBO J.* **12**: 2901-2912.

Flashner, Y. and Gralla, J. D. 1988. Dual mechanism of repression at a distance in the *lac* operon. *Proc.Natl.Acad.Sci.U.S.A.* **85**: 8968-8972.

Flavin, M. 1975 *Metabolism of sulfur compounds* Academic Press.

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, A., Tomb, J. F., Dougherty, B. A., Merrick, J. M., McKenny, K., Sutton, G. G., FitzHugh, W., Gocayne, J. D., Scott, J. L., Shirley, R., Liu, L. I., Glodek, A., Kelly, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Firtchman, J. L., Fuhrmann, J. L., Geoghagen, N. S., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and Venter, J. C. 1995. Whole genome random sequencing and assembly of *Haemophilus influenzae. Science* **269**: 496-512.

Fried, M. G. and Crothers, D. M. 1981. Equilibria and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucl.Acids Res.* **9**: 6505-6525.

Frisch, D. A., Gengenbach, B. G., Tommey, A. M., Sellner, J. M., Somers, D. A., and Myers, D. E. 1991. Isolation and characterisation of dihydropicolinate synthetase from maize. *Plant Physiol.* **96**: 444-452.

Frydman, J., Nimmesgern, E., Erdjument-Bromage, H., Wall, J. S., Tempst, P., and Hartl, F. U. 1992. Function in protein folding of TRiC, a cytosolic ring complex containing TCP-1 and structurally related subunits. *EMBO J.* **11**: 4767-4778.

Furste, J. P., Pansegrau, W., Frank, R., Blocker, H., Scholz, P., Bagdasarian, M, and Lanka, E. 1986. Molecular cloning of the plasmid RP4 primase region in a multi-host-range *tacP* expression vector. *Gene* **48**: 119-131.

Galas, D. and Schmitz, A. 1978. DNAse footprinting: a simple method for the detection of protein-DNA binding specificity. *Nucl.Acids Res.* **5**: 3157-3170.

Gao, Y., Thomas, J. O., Chow, R. L., Lee, G. H., and Cowan, N. J. 1992. A cytoplasmic chaperonin that catalyses β-actin folding. *Cell* **69**: 1043-1050.

Garner, M. M. and Revzin, A. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucl.Acids Res.* **9**: 3047-3060.

Gelb, M. H., Lin, Y., Pickard, M. A., Song, Y., and Vederas, J. C. 1990. Synthesis of 3-diaminopimelic acid isomers as inhibitors of diaminopimelate epimerase: stereocontrolled enzymatic elimination of hydrogen fluoride. *J.Am.Chem.Soc.* **112**: 4932.

Georgopoulos, C., Hendrix, R. W., Kaiser, A. D., and Wood, W. B. 1972. Role of the host cell in bacteriophage morphogenesis: effects of a bacterial mutation on T4 head assembly. *Nat.New.Biol.* **239**: 38-41.

Georgopoulos, C., Hendrix, R. W., Casjens, S. R., and Kaiser, A. D. 1973. Host participation in bacteriophage lambda head assembly. *J.Mol.Biol.* **76**: 45-60.

Georgopoulos, C. and Hohn, B. 1978. Identification of a host protein necessary for bacteriophage morphogenesis (the *groE* product). *Proc.Natl.Acad.Sci.U.S.A.* **75**: 131-135.

Georgopoulos, C. and Welch, W. J. 1993. Role of major heat shock proteins as molecular chaperones. *Annu.Rev.Cell.Biol.* **9**: 601-635.

Gerstein, M., Lesk, A. M., and Chothia, C. 1994. Structural mechanisms for domain movements in proteins. *Biochemistry* **33**: 6739-6749.

Ghislain, M., Frankard, V., and Jacobs, M. 1990. Dihydropicolinate synthetase of *Nicotiana sylvestris*, a chloroplast-localised enzyme of the lysine pathway. *Planta* **180**: 480-486.

Gilbert, W. and Maxam, A. 1973. The nucleotide sequence of the *lac* operator. *Proc.Natl.Acad.Sci.U.S.A.* **70**: 3581-3584.

Gilbert, W. 1975 *Protein-ligand interactions.* de Gruyter.

Gilvarg, C. and Bloch, K. 1951. The utilisation of acetic acid for amino acid synthesis in yeast. *J.Biol.Chem.* **193**: 339.

Gilvarg, C. 1959. N-succinyl-L-diaminopimelic acid. *J.Biol.Chem.* **234**: 2955-2959.

Gilvarg, C. 1960. Biosynthesis of diaminopimelic acid. *Fed.Proc.* **19**: 948-952.

Gilvarg, C. 1961. N-succinyl-α-amino-ε-ketopimelic acid. *J.Biol.Chem.* **234**: 2955-2959.

Girodeau, J. M., Agouridas, C., Masson, M., Pineau, R., and Le Goffic, F. 1986. The lysine pathway as a target for a new genera of synthetic antibacterial antibiotics? *J.Med.Chem.* **29**: 1023-1030.

Goenka, S. and Rao, C. M. 2001. Expression of recombinant zeta-crystallin in Escherichia coli with the help of GroEL/ES and its purification. *Protein Expr.Purif.* **21**: 260-267.

Goloubinoff, P., Gatenby, A. A., and Lorimer, G. H. 1989. GroE heat shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in *Escherichia coli. Nature* **337**: 44-47.

Gralla, J. D. 1985. Rapid "footprinting" on supercoiled DNA. *Proc.Natl.Acad.Sci.U.S.A.* **82**: 3078-3081.

Grant, S. G. N., Jessee, J., Bloom, F. R., and Hanahan, D. 1990. Differential Plasmid Rescue From Transgenic Mouse DNAs Into Escherichia coli Methylation-Restriction Mutants. *Proc.Natl.Acad.Sci.U.S.A.* **87**: 4645-4649.

Greene, R. C. 1996 *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology.* ASM Press.

Grimaud, R. and Toussaint, A. 1998. Assembly of both head and tail of bacteriophage Mu is blocked in *Escherichia coli groEL* and *groES*. *J.Bacteriol.* **180**: 1148-1153.

Grishin, N. V. and Phillips, M. A. 1994. The subunit interfaces of oligomeric enzymes are conserved to a similar extent to the overall protein sequences. *Protein Sci.* **3**: 2455-2458.

Guglielmi, G., Mazodier, P., Thompson, C. J., and Davies, J. 1991. A survey of the heat shock response in four *Streptomyces* species reveals two *groEL*-like genes and three GroEL-like proteins in *Streptomyces albus*. *J.Bacteriol.* **173**: 7374-7381.

Gupta, R. S. 1990. Sequence and structural homology between a mouse t-complex protein TCP-1 and the 'chaperonin' family of bacterial (GroEL, 60-65 kDa heat shock antigen) and eukaryotic proteins. *Biochem.Int.* **20**: 833-841.

Guyer, M. S., Reed, R. E., Steitz, T., and Low, K. B. 1981. Identification of a sex factor affinity site in *E.coli* as gamma delta. *Cold Spring Harbor Symp.Qunat.Biol.* **45**: 135-140.

Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. *J.Bacteriol.* **177**: 4121-4130.

Hahm, M. S. and Chung, B. H. 2001. Refolding and Purification of Yeast Carboxypeptidase Y Expressed as Inclusion Bodies in *Escherichia coli*. *Protein Expr.Purif.* **22**: 101-107.

Hallberg, E. M. 1990. A mitochondrial chaperonin: genetic, biochemical and molecular characteristics. *Semin.Cell Biol.* **1**: 37-45.

Hayatsu, H. and Ukita, T. 1967. The selective degradation of pyrimidines in nucleic acids by permanganate oxidation. *Biochem.Biophys.Res.Commun.* **29**: 556-561.

Haziza, C., Stragier, P., and Patte, J. C. 1982. Nucleotide sequence of the *asd* gene of *Escherichia coli*: absence of a typical attenuation signal. *EMBO J.* **1**: 379-384.

Heijenoort, J. 1996 *Escherichia coli and Salmonella: cellular and molecular biology*. ASM Press.

Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., and Dennis, D. T. 1988. Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature* **333**: 330-334.

Hendrix, R. W. and Tsui, L. 1978. Role of the host in virus assembly: cloning of the *Escherichia coli groE* gene and identification of its protein product. *Proc.Natl.Acad.Sci.U.S.A.* **75**: 136-139.

Hendrix, R. W. 1979. Purification and properties of groE, a host protein involved in bacteriophage assembly. *J.Mol.Biol.* **129**: 375-392.

Higgins, W., Tardif, C, Richaud, C., Krivanek, A., and Cardin, A. 1989. Expression of recombinant diaminopimelate epimerase in *Escherichia coli*. Isolation and inhibition with an irreversible inhibitor. *Eur.J.Biochem.* **186**: 137-143.

Hohn, T., Hohn, B., Engel, A., and Wurtz, M. 1979. Isolation and characterisation of the host protein groE involved in bacteriophage lambda assembly. *J.Mol.Biol.* **129**: 359-373.

Horovitz, A., Bochkareva, E. S., and Girshovich, A. S. 1993. The N terminus of the molecular chaperone GroEL is a crucial structural element for its assembly. *J.Biol.Chem.* **268**: 9957-9959.

Horwich, A. L., Brooks Low, K., Fenton, W. A., Hirshfield, I. N., and Furtak, K. 1993. Folding *in vivo* of bacterial cytoplasmic proteins: role of GroEL. *Cell* **74**: 909-917.

Houry, W. A., Frishman, D., Eckerskorn, C., Lottspeich, F., and Hartl, F. U. 1999. Identification of in vivo substrates of the chaperonin GroEL. *Nature* **402**: 147-154.

Höltje, J. V., Mirelman, D., Sharon, N., and Schwarz, U. 1975. Novel type of murein transglycosylase in *Escherichia coli* . *J.Bacteriol.* **124**: 1067-1076.

Hubbard, T. J. P., Ailey, B., Brenner, S. E., Murzin, A. G., and Chothia, C. 1999. SCOP: a structural classification of proteins database. *Nucl.Acids Res.* **27**: 254-256.

Huisman, O. and D'Ari, R. 1983. Effect of surpressors of SOS-mediated filamentation on *sfiA* operon expression in *Escherichia coli. J.Bacteriol.* 175.

Hutchinson, E. G., Tichelaar, W., Hofhaus, G., Weiss, H., and Leonard, K. R. 1989. Identification and electron microscopic analysis of a chaperonin oligomer from *Neurospora crassa* mitochondria. *EMBO J.* **8**: 1485-1490.

Itzhaki, L. S., Otzen, D. E., and Fersht, A. R. 1995. Nature and consequences of GroEL-protein interactions. *Biochemistry* **34**: 14581-14587.

Jacob, F. and Monod, J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J.Mol.Biol.* **3**: 318-356.

Jacob, F. and Monod, J. 1961. Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harbor Symp.Qunat.Biol.* **26**: 193.

Jaenicke, R 1991. Protein folding: local structures, domains, subunits and assemblies. *Biochemistry* **30**: 3147-3161.

Jobe, A. and Bourgeois, S. 1972. *lac* Repressor-operator interaction. VI. The natural inducer of the *lac* operon. *J.Mol.Biol.* **69**: 397-408.

Jones, E. E. and Broquist, H. P. 1966. Saccharopine an intermediate of the aminoadipic acid pathway of lysine biosynthesis. *J.Biol.Chem.* **241**: 3430-3434.

Kaneko, T., Hashimoto, T., Kumpaisal, R., and Yamada, Y. 1990. Molecular cloning of wheat dihydropicolinate synthetase. *J.Biol.Chem.* **265**: 17451-17455.

Kanemori, M., Mori, H., and Yura, T. 1994. Effects of reduced levels of GroE chaperones on protein metabolism: enhanced synthesis of heat shock proteins during steady state growth of *Escherichia coli*. *J.Bacteriol.* **176**: 4235-4242.

Kao-Huang, Y., Revzin, A., Butler, A. P., O'Conner, P., Noble, D. W., and von Hippel, P. H. 1977. Nonspecific DNA binding of genome-regulating proteins as a biological control mechanism: measurement of DNA-bound *Escherichia coli lac* repressor *in vivo*. *Proc.Natl.Acad.Sci.U.S.A.* **74**: 4228-4232.

Kelland, J. G., Palcic, M. M., Pickard, M. A., and Vederas, J. C. 1985. Stereochemistry of lysine formation by meso-diaminopimelate decarboxylase from wheat germ: use of 1H-13C NMR shift correlation to detect stereospecific deuterium labeling. *Biochemistry* **24**: 3263-3267.

Kindler, S. H. and Gilvarg, C. 1960. N-succinyl-L-α-ε-diaminopimelic acid desuccinylase. *J.Biol.Chem.* **235**: 3532-3535.

King, J., Hasse-Pettingell, C., Robinson, A. S., Speed, M., and Mitraki, A. 1996. Thermolabile folding intermediates: inclusion body precursors and chaperonin substrates. *FASEB J.* **10**: 57-66.

Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G., and Rees, D. 1996. A left-hand beta-helix revealed by the crystal structure of a carbonic anhydrase from the archaeon *Methanosarcina thermophila*. *EMBO J.* **15**: 2323-2330.

Knapp, S., Schmidt-Krey, I., Herbert, H., Bergman, T., Jornvall, H., and Ladenstein, R. 1994. The molecular chaperonin TF55 from archaeon *Sulfolobus solfataricus*. *J.Mol.Biol.* **242**: 397-407.

Kobayashi, N., Freund, S. M. V., Chatellier, J., Zahn, R., and Fersht, A. R. 1999. NMR analysis of the binding of a rhodanese peptide to a minichaperone in solution. *J.Mol.Biol.* **292**: 181-190.

Kolb, A., Busby, S., Buc, H., Garges, S., and Adhya, S. 1993. Transcriptional regulation by cAMP and its receptor protein. *Annu.Rev.Biochem.* **62**: 749-795.

Koll, H., Guiard, B., Rassow, J., Ostermann, J., Horwich, A. L., Neupert, W., and Hartl, F. U. 1992. Antifolding activity of hsp60 couples protein import into the mitochondrial matrix with export to the intermembrane space. *Cell* **68**: 1163-1175.

Kong, T. H., Coates, A. R. M., Butcher, P. D., Hickman, C. J., and Shinnick, T. M. 1993. *Mycobacterium tuberculosis* expresses two chaperonin-60 homologues. *Proc.Natl.Acad.Sci.U.S.A.* **90**: 2608-2612.

Kubota, H., Hynes, G. M., Carne, A., Ashworth, A., and Willison, K. R. 1994. Identification of six *TCP-1*-related genes encoding divergent subunits of the TCP-1 chaperonin system. *Curr.Biol.* **4**: 89-99.

Kubota, H., Hynes, G. M., and Willison, K. R. 1995. The eighth *Cct* gene, *Cctq*, encoding the theta subunit of the cytosolic chaperonin containing TCP-1. *Gene* **154**: 231-236.

Kumamoto, C. A. 1991. Molecular chaperones and protein translocation across the *Escherichia coli* inner membrane. *Mol.Microbiol.* **5**: 19-22.

Kumpaisal, R., Hashimoto, T., and Yamada, Y. 1987. Purification and characterisation of dihydropicolinate synthetase from wheat suspension cultures. *Plant Physiol.* **85**: 145-151.

Kusukawa, N. and Yura, T. 1988. Heat shock protein GroE of *Escherichia coli*. Key protective roles against thermal stress. *Genes and Dev.* **2**: 874-882.

Kusukawa, N., Yura, T., Ueguchi, C., Akiyama, Y., and Ito, K. 1989. Effects of mutations in the heat-shock genes *groES* and *groEL* on protein export of *Escherichia coli*. *EMBO J.* **8**: 3517-3521.

Laber, B., Gomis-Rüth, F. X., Romão, M., and Huber, J. 1992. *Escherichia coli* dihydropicolinate synthetase, identification of the active site and crystallisation. *Biochem.J.* **288**: 691-695.

Laemmli, V. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685.

Landry, S. J. and Gierasch, L. M. 1991. The chaperonin GroEL binds a polypeptide in an alpha-helical conformation. *Biochemistry* **30**: 7359-7362.

Landry, S. J. and Gierasch, L. M. 1991. Recognition of nascent polypeptides for targetting and folding. *Trends Biochem.Sci.* **16**: 159-163.

Landry, S. J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C., and Gierasch, L. M. 1993. Characterization of a functionally importS.ant mobile domain of GroES. *Nature* **364**: 255-258.

Langer, T., Pfeifer, G., Martin, J., Baumeister, W., and Hartl, F. U. 1992. Chaperonin-mediated protein folding: GroES binds to one end of the GroEL cylinder, which accomodates the protein substrate within its central cavity. *EMBO J.* **11**: 4757-4765.

Laskey, R. A. and Earnshaw, W. C. 1980. Nucleosome assembly. *Nature* **286**: 763-767.

Lawther, R. P., Calhoun, D. H., Adams, C. W., Hauser, C. A., Gray, J., and Hatfield, G. W. 1981. Molecular basis of valine resistance in *Escherichia coli* K-12. *Proc.Natl.Acad.Sci.U.S.A.* **78**: 922-925.

Ledwidge, R. and Blanchard, J. S. 1999. The dual biosynthetic capability of *N*-acetylornithine aminotransferase in argenin and lysine biosynthesis. *Biochemistry* **38**: 3019-3024.

Lee, J. and Goldfarb, A. 1991. *lac* repressor acts by modifying the initial transcribing complex so that it cannot leave the promoter. *Cell* **66**: 793-798.

Lee, S. C. and Olins, P. O. 1992. Effect of overproduction of heat shock chaperones GroESL and DnaK on human procollagenase production in *Escherichia coli*. *J.Biol.Chem.* **267**: 2849-2852.

Lehel, C., Los, D., Wada, H., Gyorgyei, J., Horvath, I., Kovacks, E., Murata, N., and Vigh, L. 1993. A second *groEL*-like gene, organised in a *groESL* operon, is present in the genome of *Synechocystis* sp. PCC-6803. *J.Biol.Chem.* **268**: 1901-1907.

Lenstra, J. A. and Bloemendal, H. 1983. Topography of the total protein population from cultured cells upon fractionation by chemical extraction. *Eur.J.Biochem.* **135**: 413-423.

Lewis, M., Chang, J., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G., and Lu, P. 1996. Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* **271**: 1247-1254.

Lewis, V. A., Hynes, G. M., Zheng, D., Saibil, H. R., and Willison, K. R. 1992. T-complex polypeptide-1 is a subunit of a heteromeric particle in the eukaryotic cytosol. *Nature* **358**: 249-252.

Lin, S and Riggs, A. D. 1975. A comparison of *lac* repressor binding to operator and non-operator DNA. *Biochem.Biophys.Res.Commun.* **62**: 704-710.

Lin, S and Riggs, A. D. 1975. The general affinity of *lac* repressor for *E. coli* DNA: implications for gene regulation in procaryotes and eucaryotes. *Cell* **4**: 107-111.

Lin, Y., Myhrman, R., Schrag, M. L., and Gelb, M. H. 1988. Bacterial N-succinyl-L-diaminopimelic acid desuccinylase. Purification, partial characterisation, and substrate specificity. *J.Biol.Chem.* **263**: 1622-1627.

Lin, Z., Schwartz, F. P., and Eisenstein, E. 1995. The hydrophobic nature of GroEL-substrate binding. *J.Biol.Chem.* **270**: 1011-1014.

Lingappa, J. R., Martin, R. L., Wong, M. L., Gamen, D., Welch, W. J., and Lingappa, V. R. 1994. A eukaryotic cytosolic chaperonin is associated with a high molecular mass intermediate in the assembly of hepatitis B virus capsid, a multimeric particle. *J.Cell Biol.* **125**: 99-111.

Link, A. J., Phillips, D., and Church, G. M. 1997. Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterisation. *J.Bacteriol.* **179**: 6228-6237.

Lissin, N. M., Venyaminov, S. Y., and Girshovich, A. S. 1990. (Mg-ATP)-dependent self assembly of molecular chaperone GroEL. *Nature* **348**: 339-342.

Lobel, R. B. and Schleif, R. F. 1990. DNA looping and unlooping by AraC protein. *Science* **250**: 528-532.

Lorimer, G. H. 1996. A quantitative assessment of the role of the chaperonin proteins in protein folding *in vivo*. *FASEB J.* **10**: 5-9.

Lorimer, G. H. 1997. Protein folding. Folding with a two stroke motor. *Nature* **388**: 720-721.

Losick, R. and Gilvarg, C. 1966. Effect of α-acetylation on utilisation of lysine oligopeptides in *E.coli*. *J.Biol.Chem.* **240**: 2340-2346.

Lowen, P. C. and Hengge-Aronis, R. 1994. The role of sigma factor σ^S (*katF*) in global regulation. *Annu.Rev.Microbiol.* **48**: 53-80.

Lubben, T. H., Donaldson, G. K., Viitanen, P. V., and Gatenby, A. A. 1989. Several proteins imported into the chloroplast form stable complexes with the GroEL-related chloroplast molecular chaperone. *The Plant Cell* **1**: 1223-1230.

Lutz, R. and Bujard, H. 1997. Independent and tight regulation of transcriptional untis in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I_1-I_2 regulatory elements. *Nucl.Acids Res.* **25**: 1203-1210.

Macario, A. J. L., Lange, M., Ahring, B. K., and De Macario, E. C. 1999. Stress genes and proteins in the archaea. *Microbiol.Mol.Biol.Rev* **63**: 923-967.

Maniatis, T., Fritsch, E. F., and Sambrook, J. 1983. *Molecular Cloning: A laboratory manual.* Cold Spring Harbor Laboratory.

Martin, C., Cami, B., Yeh, P., Stragier, P., Parsot, C., and Patte, J. C. 1988. *Pseudomonas aeruginosa* diaminopimelate decarboxylase: evolutionary relationship with other amino acid decarboxylases. *Mol.Biol.Evol.* **5**: 549-559.

Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A. L., and Hartl, F. U. 1991. Chaperonin-mediated protein folding at the surface of GroEL through a 'molten globule' like intermediate. *Nature* **352**: 36-42.

Matthews, C. R. 1993. Pathways of protein folding. *Annu.Rev.Biochem.* **62**: 653-683.

Matthews, K. S. and Nichols, J. C. 1998. Lactose repressor protein: functional properties and structure. *Prog.Nucl.Acid Res.* **58**: 127-164.

Maxam, A. and Gilbert, W. 1977. A new method for sequencing DNA. *Proc.Natl.Acad.Sci.U.S.A.* **74**: 560-564.

McLennan, N., Girshovich, A. S., Lissin, N. M., Charters, Y., and Masters, M. 1993. The strongly conserved carboxyl-terminus glycine-methionine motif of the *Escherichia coli* GroEL cahperonin is dispensable. *Mol.Microbiol.* **7**: 49-58.

McLennan, N. and Masters, M. 1998. GroE is vital for cell wall synthesis. *Nature* **392**: 139.

McMullin, T. W. and Hallberg, R. L. 1988. A highly evolutionarily conserved mitochondrial protein is structurally related to the protein encoded by the *Escherichia coli gorEL* gene. *Mol.Cell.Biol.* **8**: 371-380.

Meinnel, T., Schmitt, E., Mechulam, Y., and Blanquet, S. 1992. Structural and biochemical characterization of the *Escherichia coli argE* gene product. *J.Bacteriol.* **174**: 2323-2331.

Mengin-Lecreulx, D., Blanot, D., and Heijenoort, J. 1994. Replacement of diaminopimelic acid by cystathionine of lanthionine in the peptidoglycan of *Escherichia coli*. *J.Bacteriol.* **176**: 4321-4327.

Miller, J. H. 1972 *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory, New York.

Mirwaldt, C., Korndorfer, I., and Huber, R. 1995. The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 A resolution. *J.Mol.Biol.* **246**: 227-239.

Misono, H., Togawa, H., Yamamoto, T., and Soda, K. 1976. Occurrence of meso-alpha, epsilon-diaminopimelate dehydrogenase in *Bacillus sphaericus*. *Biochem.Biophys.Res.Commun.* **72**: 89-93.

Misono, H., Togawa, H., Yamamoto, T., and Soda, K. 1979. Meso-alpha,epsilon-diaminopimelate D-dehydrogenase: distribution and the reaction product. *J.Bacteriol.* **137**: 22-27.

Misono, H., Ogasawara, M., and Nagasaki, S. 1986. Identification of D-L-diaminopimelate dehydrogenase in *Corynebacterium glutamicum*. *Biol.Chem.* **50**: 1329-1330.

Morett, E. and Buck, M. 1988. NifA-dependent in vivo protection demonstrates that the upstream activator sequence of *nif* promoters is a protein binding site. *Proc.Natl.Acad.Sci.U.S.A.* **85**: 9401-9405.

Murayama, N., Shimizu, H., Takiguchi, S., Baba, Y., Amino, H., Horiuchi, T., Sekimizu, K., and Miki, T. 1996. Evidence for involvement of *Escherichia coli* genes *pmbA*, *cspA* and a previously unrecognised gene *tldD*, in the control of DNA gyrase by *letD* (*ccdB*) of sex factor F. *J.Mol.Biol.* **256**: 483-502.

Nelson, M. and McClelland, M. 1992. Use of DNA methyltransferase/endonuclease enzyme combinations for megabase mapping of chromosomes. *Methods Enzymol.* **216**: 279-303.

Nielsen, K. L. and Cowan, N. J. 1998. A single-ring is sufficient for productive chaperonin-mediated folding *in vivo*. *Mol.Cell* **2**: 1-7.

Nielsen, K. L., McLennan, N., Masters, M., and Cowan, N. J. 1999. A single-ring mitochondrial chaperonin (Hsp60-Hsp10) can substitute for GroEL-GroES *in vivo*. *J.Bacteriol.* **181**: 5871-5875.

O'Farrell, P. H. 1975. High resolution two-dimensional electrophoresis of proteins. *J.Biol.Chem.* **250**: 4007-4021.

Oehler, S., Eismann, E. R., Kramer, H., and Muller-Hill, B. 1990. The three operators of the *lac* operon cooperate in repression. *EMBO J.* **9**: 973-979.

Ogata, R. and Gilbert, W. 1979. DNA binding site of *lac* repressor probed by dimthhylsulfate methylation of *lac* operator. *J.Mol.Biol.* **132**: 709-728.

Old, I. G., Saint-Girons, I., and Richaud, F. C. 1991. Regulation of methionine biosynthesis in the *Enterobacteriaceae*. *Prog.Biophys.Mol.Biol.* **56**: 145-185.

Orengo, C. A. 1999. THE CATH database provides insights into protein structure/function relationships. *Nucl.Acids Res.* **27**: 275-279.

Osterman, A. L., Kinch, L. N., Grishin, N. V., and Phillips, M. A. 1995. Acidic residues important for substrate binding and cofactor reactivity in eukaryotic ornithine decarboxylase identified by alanine scanning mutagenesis. *J.Biol.Chem.* **270**: 11797-11802.

Panse, V. G., Udgaonkar, J. B., and Varadarajan, R. 1998. SecB binds only to a late native-like intermediate in the folding pathway of barstar and not to the unfolded state. *Biochemistry* **37**: 14477-14483.

Patte, J. C., Loviny, T., and Cohen, G. N. 1962. Repression de la décarboxylase de l'acide meso αε-diaminopimelique par la lysine chez *E.coli*. *Biochim.Biophys.Acta.* **58**: 359-360.

Patte, J. C., Loviny, T., and Cohen, G. N. 1965. Effets inhibiteurs cooperatifs de la L-lysine avec d'autres amino acides sur une aspartokinase d'*Escherichia coli*. *Biochim.Biophys.Acta.* **99**: 523-530.

Patte, J. C., Lebras, G., and Cohen, G. N. 1967. Regulation by methionine of the synthesis of a third aspartokinase and a second homoserine dehydrogenase in *Escherichia coli* K12. *Biochim.Biophys.Acta.* **136**: 245-257.

Patte, J. C., Morand, P., Boy, E., Richaud, C., and Borne, F. 1980. The *relA* locus and the regulation of lysine biosynthesis in *Escherichia coli*. *Mol.Gen.Genet.* **179**: 319-325.

Patte, J. C. 1983 *Amino acid biosynthesis and genetic regulation* Addison-Wesley Publishing Co.

Patte, J. C. 1996 *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology.* ASM Press.

Pavelka, M. S. and Jacobs, W. R. 1996. Biosynthesis of diaminopimelate, the precursor of lysine and a component of peptidoglycan, is an essential function of *Mycobacterium smegmatis*. *J.Bacteriol.* **178**: 6496-6507.

Pavelka, M. S., Weisbrod, T. R., and Jacobs, W. R. 1997. Cloning of the dapB gene, encoding dihydrodipicolinate reductase, from *Mycobacterium tuberculosis*. *J.Bacteriol.* **179**: 2777-2782.

Perl, A., Shaul, O., and Galili, G. 1992. Regulation of lysine synthesis in transgenic potato plants expressing a bacterial dihydropicolinate synthetase in their chloroplasts. *Plant Mol.Biol.* **19**: 815-823.

Perrett, S., Zahn, R., Sternberg, N., and Fersht, A. R. 1997. Importance of electrostatic interactions in the rapid binding of polypeptides to GroEL. *J.Mol.Biol.* **269**: 892-901.

Peterkofsky, B. and Gilvarg, C. 1961. N-succinyl-L-diaminopimelic-glutamic trasaminase. *J.Biol.Chem.* **236**: 1432-1438.

Peterkofsky, B. 1962. *N*-succinyl-L-diaminopimelate glutamic transaminase. *Methods Enzymol.* **5**: 853-858.

Peters, J. M. 1994. Proteasomes: protein degradation machines of the cell. *Trends Biochem.Sci.* **19** : 377-382.

Phipps, B. M., Typke, D., Hegerl, R., Volker, S., Hoffmann, A., Stetter, K. O., and Baumeister, W. 1993. Structure of a molecular chaperone from a thermophilic archaebacterium. *Nature* **361**: 475-477.

Ptitsyn, O. B. 1995. Structures of folding intermediates. *Curr.Opin.Struct.Biol.* **5**: 740-778.

Radford, S. E., Dobson, C. M., and Evans, P. A. 1992. The folding of hen egg white lysozyme: a complex process involving partially structured intermediates and multiple pathways. *Nature* **358**: 302-307.

Raetz, C. R. H. and Roderick, S. L. 1995. A left-handed parallel beta helix in the structure of UDP-N-acetylglucosamine acyltransferase. *Science* **270**: 997-1000.

Ranson, N., White, H., and Saibil, H. R. 1998. Chaperonins. *Biochem.J.* **333**: 233-242.

Reading, D. S., Hallberg, R. L., and Myers, A. M. 1989. Characterisation of a yeast HSP60 gene coding for a mitochondrial assembly factor. *Nature* **337**: 655-659.

Reeder, T. and Schleif, R. F. 1993. AraC protein can activate transcription from only one position and when pointed in only one direction. *J.Mol.Biol.* **231**: 205-218.

Revel, H. R., Stitt, B. L., Lielausis, I., and Wood, W. B. 1980. Role of the host cell in bacteriophage T4 development. *J.Virol.* **33**: 366-376.

Richaud, C., Higgins, W., Mengin-Lecreulx, D., and Stragier, P. 1987. Molecular cloning, characterisation, and chromosomal location of *dapF*, the *Escherichia coli* gene for diaminopimelate epimerase. *J.Bacteriol.* **169**: 1454-1459.

Richaud, F., Richaud, F. C., Ratet, P., and Patte, J. C. 1986. Chromosomal location and nucleotide sequence of the *Escherichia coli dapA* gene. *J.Bacteriol.* **166**: 297-300.

Richaud, F. C., Richaud, F., Haziza, C., and Patte, J. C. 1984. Regulation of expression and nucleotide sequence of the *E.coli dapD* gene. *J.Biol.Chem.* **259**: 14824-14829.

Richaud, F. C. and Printz, C. 1988. Nucleotide sequence of the *dapF* gene and flanking regions from *Escherichia coli* K12. *Nucl.Acids Res.* **16**: 10367.

Richaud, F. C., Mengin-Lecreulx, D., Pochet, S., Johnson, E. J., Cohen, G. N., and Marlière, P. 1993. Directed evolution of biosynthetic pathways: recruitment of cysteine thioethers for constructing the cell wall of *Escherichia coli*. *J.Biol.Chem.* **268**: 26827-26835.

Riggs, A. D., Suzuki, H., and Bourgeois, S. 1970. Lac repressor-operator interaction. I. Equilibrium studies. *J.Mol.Biol.* **48**: 67-83.

Rommelaere, H., van Troys, M., Gao, Y., Melki, R., Cowan, N. J., Van-dekerckhove, J., and Ampe, C. 1993. Eukaryotic cytosolic chaperonin contains t-complex polypeptide-1 and seven related subunits. *Proc.Natl.Acad.Sci.U.S.A.* **90**: 11975-11979.

Roobol, A. and Carden, M. J. 1993. Identification of chaperonin particles in mammalian brain cytosol and t-complex polypeptide-1 as one of their components. *J.Neurochem.* **60**: 2327-2330.

Roobol, A., Holmes, F. E., Hayes, N. V. L., Baines, A. J., and Carden, A. 1995. Cytoplasmic chaperonin complex enters neurites developing *in vitro* and differ in subunit composition within single cells. *J.Cell Sci.* **108**: 1477-1488.

Roseman, A. M., Chen, S., White, H., Braig, K., and Saibil, H. R. 1996. The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate binding domains in GroEL. *Cell* **87**: 241-251.

Rosenberg, H. F., Ackerman, S. J., and Tenen, D. G. 1993. Characterisation of a distinct binding site for the prokaryotic chaperone, GroEL, on a human granulocyte ribonuclease. *J.Biol.Chem.* **268**: 4499-4503.

Rowsell, S., Pauptit, R. A., Tucker, A. D., Melton, R. G., Blow, D. M., and Brick, P. 1997. Crystal structure of carboxypeptidase G2, a bacterial enzyme with applications in cancer therapy. *Structure* **5**: 337-347.

Rusanganwa, E. and Gupta, R. S. 1993. Cloning and characterisation of multiple *groEL* chaperonin-encoded genes in *Rhizobium meliloti*. *Gene* **126**: 67-75.

Rye, H. S., Burston, S. G., Fenton, W. A., Beechem, J. M., Xu, Z., Sigler, P. B., and Horwich, A. L. 1997. Distinct actions of cis and trans ATP within the double ring of the chaperonin GroEL. *Nature* **388**: 792-798.

Sagisaka, S. and Shimura, K. 1961. Studies on lysine biosynthesis II. Metabolic fate of DL-α-aminoadipic acid 6-^{14}C in *Torulopsis utilis*. *J.Biochem.* **49**: 392.

Saibil, H. R., Zheng, D., Roseman, A. M., Hunter, A. S., Watson, G. M. F., Chen, S., auf der Mauer, A., O'Hara, B. P., Wood, S. P., Mann, N. H., Barnett, L. K., and Ellis, R. J. 1993. ATP induces large quaternary rearrangements in a cage like chaperonin structure. *Curr.Biol.* **3**: 265-273.

Saint-Girons, I., Parsot, C., Zakin, M. M., Barzu, O., and Cohen, G. N. 1988. Methionine biosynthesis in *Enterobacteriaceae* : biochemicsal, regulatory and evolutionary aspects. *Crit.Rev.Biochem.* **23**: S1-S42.

Sakikawa, C., Taguchi, H., Makino, Y., and Yoshida, M. 1999. On the maximum size of proteins to stay and fold in the cavity of GroEL underneath GroES. *J.Biol.Chem.* **274**: 21251-21256.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain terminating inhibitors. *Proc.Natl.Acad.Sci.U.S.A.* **74**: 5463-5467.

Sasse-Dwight, S. and Gralla, J. D. 1989. KMnO4 as a probe for *lac* promoter DNA melting and mechanism in vivo. *J.Biol.Chem.* **264**: 8074-8081.

Sasse-Dwight, S. and Gralla, J. D. 1990. Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor sigma 54. *Cell* **62**: 945-954.

Saunders, P. P. and Broquist, H. P. 1966. Saccharopine an intermediate of the aminoadipic acid pathway of lysine biosynthesis. IV. Saccharopine dehydrogenase. *J.Biol.Chem.* **241**: 3435-3440.

Scapin, G., Blanchard, J. S., and Sacchettini, J. C. 1995. Three-dimensional structure of Escherichia coli dihydrodipicolinate reductase. *Biochemistry* **34**: 3502-3512.

Scapin, G. and Blanchard, J. S. 1998. Enzymology of bacterial lysine biosynthesis. *Adv.Enzymol.Relat.Areas Mol.Biol.* **72**: 279-323.

Schlax, P. J., Capp, W., and Record, M. T. 1995. Inhibition of transcription initiation by *lac* repressor. *J.Mol.Biol.* **245**: 331-350.

Schleif, R. F. 1992. DNA looping. *Annu.Rev.Biochem.* **61**: 199-223.

Schmidt, M. and Buchner, J. 1992. Interaction of GroEL with an all-β protein. *J.Biol.Chem.* **267**: 16829-16833.

Schrumpf, B., Schwarzer, A., Kalinowski, J., Puhler, A., Eggeling, L., and Sahm, M. 1991. A functionally split pathway for lysine synthesis in *Corynebacterium glutamicium*. *J.Bacteriol.* **173**: 4510.

Shalloway, D., Kleinberger, T., and Livingston, D. M. 1980. Mapping of SV40 DNA replication origin region binding sites for the SV40 T antigen by protection against exonuclease III digestion. *Cell* **20**: 411-422.

Shaul, O. and Galili, G. 1993. Concerted regulation of lysine anthreonine synthesis in tobacco plants expressing feedback insensitive aspartate kinase and dihydropicolinate synthetase. *Plant Mol.Biol.* **23**: 759-768.

Shaver, J. M., Bittel, D., Sellner, J. M., Frisch, D. A., Somers, D. A., and Gengenbach, B. G. 1996. Single-amino acid substitutions eliminate lysine inhibition of maize dihydrodipicolinate synthase. *Proc.Natl.Acad.Sci.U.S.A.* **93**: 1962-1966.

Shaw, J. P., Petsko, G., and Ringe, D. 1997. Determination of the structure of alanine racemase from *Bacillus stearothermophilus* at 1.9-A resolution. *Biochemistry* **36**: 1329-1342.

Shedlarsky, J. G. and Gilvarg, C. 1970. The pyruvate-aspartic semialdehyde condensing enzyme of *Escherichia coli*. *J.Biol.Chem.* **245**: 1362-1373.

Shtilerman, M., Lorimer, G. H., and Englander, S. W. 1999. Chaperonin function: folding by forced unfolding. *Science* **284**: 822-825.

Simms, S. A., Voige, W. H., and Gilvarg, C. 1984. Purification and characterization of succinyl-CoA: tetrahydrodipicolinate N-succinyltransferase from *Escherichia coli*. *J.Biol.Chem.* **259**: 2734-2741.

Simon, L. D., Snover, D., and Doermann, A. H. 1974. Bacterial mutation affecting T4 phage DNA synthesis and tail production. *Nature* **252**: 451-455.

Simons, R. W., Houman, F., and Kleckner, N. 1987. Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene* **53**: 85-96.

Singer, M., Baker, T. A., Schnitzler, G., Deischel, S. M., Goel, M., Dove, W. F., Jaacks, K. J., Grossman, A. D., Erickson, J., and Gross, C. A. 1989. A collection of strains containing genetically linked alternating antibiotic resistance elements for genetic mapping of *Escherichia coli*. *Microbiol.Rev.* **53**: 1-24.

Song, Y., Niederer, D., Lane-Bell, P. M., Lam, L. K. P., Crawely, S., Palcic, M. M., Pickard, M. A., Preuss, D. L., and Vederas, J. C. 1994. Stereospecific analysis of phosphonate analogues of diaminopimelic acid (DAP), their interaction with DAP enzymes, and antibacterial activity of peptide derivatives. *J.Org.Chem.* **59**: 5784-5793.

Stadtman, E. R., Cohen, G. N., Lebras, G., and de Robichon-Szulmajster, H. 1961. Feedback inhibition and repression of aspartokinase activity in *E.coli* and *Saccharomyces cerevisiae*. *J.Biol.Chem.* **236**: 2033-2038.

Sternberg, N. 1973. Properties of a mutant of *Escherichia coli* defective in bacteriophage λ head formation (*groE*) II. *J.Mol.Biol.* **76**: 25-44.

Sternberg, N. 1973. Properties of a mutant of *Escherichia coli* defective in bacteriophage λ head formation (*groE*) I. Initial characterisation. *J.Mol.Biol.* **76**: 1-23.

Sternlicht, H., Farr, G., Sternlicht, M. L., Driscoll, J. K., Willison, K. R., and Yaffe, M. B. 1993. The t-complex polypeptide-1 complex is a chaperonin for tubulin and actin *in vivo*. *Proc.Natl.Acad.Sci.U.S.A.* **90**: 9422-9426.

Stragier, P., Danos, O., and Patte, J. C. 1983. Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli*. II. Nucleotide sequence of the *lysA* gene and its regulatory region. *J.Mol.Biol.* **168**: 321-331.

Stragier, P., Richaud, F., Borne, F., and Patte, J. C. 1983. Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli* I. Identification of a *lysR* gene encoding an activator of the *lysA* gene. *J.Mol.Biol.* **168**: 307-320.

Stragier, P. and Patte, J. C. 1983. Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli* III. Nucleotide sequence and regulation of the *lysR* gene. *J.Mol.Biol.* **168**: 333-350.

Straney, S. B. and Crothers, D. M. 1987. Lac repressor is a transient gene-activating protein. *Cell* **51**: 699-707.

Strassman, M. and Weinhouse, S. 1953. Biosynthetic pathways III. The biosynthesis of lysine by *Torulopsis utilis*. *J.Am.Chem.Soc.* **75**: 1680.

Strassman, M. and Ceci, L. N. 1964. Enzymatic formation of homcitric acid, an intermediate in lysine biosynthesis. *Biochem.Biophys.Res.Commun.* **14**: 262-267.

Strassman, M. and Ceci, L. N. 1965. Enzymatic formation of α-ketoadipic acid from homoisocitric acid. *J.Biol.Chem.* **240**: 4357-4361.

Strassman, M. and Ceci, L. N. 1966. Enzymatic formation of *cis*-homoaconitic acid, an intermediate in lysine biosynthesis in yeast. *J.Biol.Chem.* **241**: 5401-5407.

Studier, F. W. and Moffat, B. A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high level expression of cloned genes. *J.Mol.Biol.* **189**: 113-130.

Sundharadas, G. and Gilvarg, C. 1967. Biosynthesis of alpha,epsilon-diaminopimelic acid in *Bacillus megaterium*. *J.Biol.Chem.* **242**: 3983-3988.

Surewicz, W. K., Mantsch, H. H., and Chapman, D. 1993. Determination of protein secondary structure by Fourier transform infrared spectroscopy: a critical assessment. *Biochemistry* **32**: 389-394.

Swaisgood, H. E. 1992 *Advanced dairy chemistry 1: Proteins* Elsevier science publishers, LTD.

Tabor, S. and Richardson, C. C. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc.Natl.Acad.Sci.U.S.A.* **82**: 1074-1078.

Tabor, S. and Richardson, C. C. 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc.Natl.Acad.Sci.U.S.A.* **84**: 4767-4771.

Taguchi, H., Makino, Y., and Yoshida, M. 1994. Monomeric chaperonin-60 and its 50-kDa fragment possess the ability to interact with non-native proteins, to suppress aggregation and to promote protein folding. *J.Biol.Chem.* **269**: 8529-8534.

Takahashi, H., Coppo, A., Manzi, A., Martire, G., and Pulitzer, J. F. 1975. Design of a system of conditional lethal mutations (tab/k/com) affecting protein-protein interactions in bacteriophage T4-infected *Escherichia coli*. *J.Mol.Biol.* **96**: 563-578.

Takano, T. and Kakefuda, T. 1972. Involvement of a bacterial factor in morphogenesis of bacteriophage capsid. *Nat.New.Biol.* **239**: 34-37.

Tamir, H. and Gilvarg, C. 1974. Dihydropicolinic acid reductase. *J.Biol.Chem.* **249**: 3034-3040.

Tanaka, N. and Fersht, A. R. 1999. Identification of substrate binding site of GroEL minichaperone in solution. *J.Mol.Biol.* **292**: 173-180.

Tilly, K., Murialdo, H., and Georgopoulos, C. 1981. Identification of a second *Escherichia coli groE* gene whose product is necessary for bacteriophage morphogenesis. *Proc.Natl.Acad.Sci.U.S.A.* **78**: 1629-1633.

Trent, J. D., Nimmesgern, E., Wall, J. S., Hartl, F. U., and Horwich, A. L. 1991. A molecular chaperone froma thermophilic archaebacterium is related to the eukaryotic protein t-complex polypeptide-1. *Nature* **354**: 490-493.

Truffa-Bachi, P., Patte, J. C., and Cohen, G. N. 1967. Sur la dihydropicolinate synthétase d'*Escherichia coli* K12. *Comptes Rendus Acad.Sci.Paris* **265**: 928-929.

Truffa-Bachi, P., van Rapenbusch, R., Janin, J., Gross, C. A., and Cohen, G. N. 1968. The threonine-sensitive homoserine dehydrogenase and aspartokinase activities of *Escherichia coli* K12. 4. Isolation, molecular weight, amino acid analysis and behaviour of the sulfhydryl groups of the protein catalysing the two activities. *Eur.J.Biochem.* **5**: 73-80.

Umbarger, H. E. 1996 *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology.* ASM Press.

Vaara, M. 1992. Eight bacterial proteins, including UDP-N-acetylglucosamine acyltransferase (LpxA) and three other transferases of *Escherichia coli*, consist of a six-residue periodicity theme. *FEMS Microbiol.Letts.* **97**: 249-254.

van Stokkum, I. H., Spoelder, H. J., Bloemendal, M., van Grondelle, R., and Groen, F. C. 1990. Estimation of protein secondary structure and error analysis from circular dichroism spectra. *Anal.Biochem.* **191**: 110-118.

Vicente, M., Gomez, M. J., and Ayala, J. A. 1998. Regulation of transcription of cell division genes in the *Escherichia coli* dcw cluster. *Cell Mol.Life Sci.* **54**: 317-324.

Viitanen, P. V., Lubben, T. H., Reed, J., Goloubinoff, P., O'Keefe, D., and Lorimer, G. H. 1990. Chaperonin-facilitated refolding of ribulosebisphosphate carboxylase and ATP hydrolysis by chaperonin 60 (*groEL*) are K+ dependent. *Biochemistry* **29**: 5665-5671.

Viitanen, P. V., Gatenby, A. A., and Lorimer, G. H. 1992. Purified chaperonin 60 (GroEL) interacts with the non-native states of a multitude of *Escherichia coli* proteins. *Protein Sci.* **1**: 363-369.

von Hippel, P. H., Revzin, A., Gross, C. A., and Wang, A. C. 1974. Non-specific DNA binding of genome regulating proteins as a biological control mechanism: I. The *lac* operon: equilibrium aspects. *Proc.Natl.Acad.Sci.U.S.A.* **71**: 4808-4812.

Wallington, E. J. and Lund, P. A. 1994. *Rhizobium leguminosarum* contains multiple chaperonin (*cpn60*) genes. *Microbiology* **140**: 113-122.

Wang, J. D. and Weissman, J. S. 1999. Thinking outside the box: new insights into the mechanism of GroEL-mediated protein folding. *Nature Structural Biology* **6**: 597-600.

Wang, M. M. and Reed, R. R. 1993. Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast *Nature* **364**: 121-126.

Wang, Q, Buckle, A. M., and Fersht, A. R. 2000. From minichaperone to GroEL 1: Information on GroEL-polypeptide interations from crystal packing of minichaperones. *J.Mol.Biol.* **304**: 873-881.

Weber, F., Keppel, F., Georgopoulos, C., Hayer-Hartl, M. K., and Hartl, F. U. 1998. The oligomeric structure of GroEL/GroES is required for biologically significant chaperonin function in protein folding. *Nature Structural Biology* **5**: 977-985.

Weissman, J. S., Hohl, C. M., Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H. R., Fenton, W. A., and Horwich, A. L. 1995. Mechanism of GroEL action: productive release of polypeptide from a sequestered position under GroES. *Cell* **83**: 577-587.

Weissman, J. S., Rye, H. S., Fenton, W. A., Beechem, J. M., and Horwich, A. L. 1996. Characterisation of the active intermediate of a GroEL-GroES-mediated protein folding reaction. *Cell* **84**: 481-490.

White, P. J. and Kelly, B. 1965. Purification and properties of diaminopimelate decarboxylase from *Escherichia coli*. *Biochem.J.* **96**: 75-84.

White, P. J. 1983. The essential role of diaminopimelate dehydrogenase in the biosynthesis of lysine by *Bacillus sphaericus* . *J.Gen.Microbiol.* **129**: 739-749.

Whittaker, P. A., Campbell, A. J. B., Southern, E. M., and Murray, N. E. 1988. Enhanced recovery and restriction mapping of DNA fragments cloned into a new λ vector. *Nucl.Acids Res.* **16**: 6725-6736.

Wick, K. L. and Matthews, K. S. 1991. Interactions between *lac* repressor protein and site-specific bromodeoxyuridine-substituted operator DNA. Ultraviolet footprinting and protein-DNA cross-link formation. *J.Biol.Chem.* **266**: 6106-6112.

Willison, K. R. and Kubota, H. 1994 *The biology of heat shock proteins and molecular chaperones.* CSH Laboratory Press.

Wiseman, J. S. and Nichols, J. S. 1948. Purification and properties of diaminopimelic acid epimerase from *Escherichia coli. J.Biol.Chem.* **259**: 8907-8914.

Work, E. 1950. A new naturally occurring amino acid. *Nature* **165**: 74-75.

Work, E. 1951. The isolation of αε-diaminopimelic acid from *Corynebacterium diptheriae* and *Mycobacterium tuberculosis. Biochem.J.* **49**: 17-23.

Wu, W. F., Urbanowski, M. L., and Stauffer, G. V. 1993. Role of the MetR regulatory system in vitamin B_{12}-mediated repression of the *Salmonella typhimurium metE* gene. *J.Bacteriol.* **174**: 4833-4837.

Xu, Z., Horwich, A. L., and Sigler, P. B. 1997. The crystal structure of the assymetric GroEL-GroES-$(ADP)_7$ chaperonin complex. *Nature* **388**: 741-750.

Yamamoto, J., Shimizu, M., and Yamane, K. 1989. Nucleotide sequence of the diaminopimelate-decarboxylase gene from *Bacillus subtilis. Nucl.Acids Res.* **17**: 10105.

Yanisch-Perron, C., Vieira, J., and Messing, J. I. 1985. Improved M13 phage cloning vecotrs and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-109.

Yaphe, W. 1957. The use of agarase from *Pseudomonas atlantica* in the identification of agar in marine algae (*Rhodophyceae*). *Can.J.Microbiol.* **3**: 987-993.

Yeh, P., Sicard, A. M., and Sinskey, A. J. 1988. Nucleotide sequence of the *lysA* gene of *Corynebacterium glutamicum* and possible mechanisms for modulation of its expression. *Mol.Gen.Genet.* **212**: 112-119.

Yifrach, O. and Horovitz, A. 1994. Two lines of allosteric communication in the oligomeric chaperonin GroEL are revealed by the single mutation Arg196®Ala. *J.Mol.Biol.* **243**: 397-401.

Yifrach, O. and Horovitz, A. 1998. Transient kinetic analysis of adenosine 5'-triphosphate bindign-induced conformational chnages in the allosteric chaperonin GroEL. *Biochemistry* **37**: 7083-7088.

Yifrach, O. and Horovitz, A. 2000. Coupling between protein folding and allostery in the GroE chaperonin system. *Proc.Natl.Acad.Sci.U.S.A.* **97**: 1521-1524.

Yugari, Y. and Gilvarg, C. 1965. The condensation step in diaminopimelate synthesis. *J.Biol.Chem.* **240**: 4710-4716.

Yung, H. L., Heller, K., Gellert, M., and Zubay, G. 1979. Differential sensitivity of gene expression *in vitro* to inhibition of DNA gyrase. *Proc.Natl.Acad.Sci.U.S.A.* **76**: 3304-3308.

Zahn, R. and Plückthun, A. 1992. GroE prevents the accumulation of early folding intermediates of pre-β-lactamase without changing the folding pathway. *Biochem.* **31**: 3249-3255.

Zahn, R., Buckle, A. M., Perrett, S., Johnson, C. M. J., Corrales, F. J., Golbik, R., and Fersht, A. R. 1996. Chaperone activity and structure of monomeric polypeptide binding domains of GroEL. *Proc.Natl.Acad.Sci.U.S.A.* **93**: 15024-15029.

Zhang, L. and Gralla, J. D. 1989. Micrococcal nuclease as a probe for bound and distorted DNA in lac transcription and repression complexes. *Nucl.Acids Res.* **17**: 5017-5028.

Zhang, L. and Gralla, J. D. 1990. In situ nucleoprotein structure involving origin-proximal SV40 DNA control elements. *Nucl.Acids Res.* **18**: 1797-1803.

Zubay, G., Schwartz, D., and Beckwith, J. 1970. Mechanism of activation of catabolite-sensitive genes: a positive control system. *Proc.Natl.Acad.Sci.U.S.A.* **66**: 104-110.

Zweig, M. and Cummings, D. J. 1973. Cleavage of head and tail proteins during bacteriophage T5 assembly: selective host involvement in the cleavage of a tail protein. *J.Mol.Biol.* **80**: 505-518.

Zwickl, P., Pfeifer, G., Lottspeich, F., Kopp, F., Dahlmann, B., and Baumeister, W. 1990. Electron microscopy and image analysis reveal common principles of organisation in two large protein complexes: *groEL*-type proteins and proteasomes. *J.Struct.Biol.* **103** : 197-203.